DOCUMENT RESUME

ED 481 263 CS 512 488

TITLE	Proceedings of the Annual Meeting of the Association for Education in Journalism and Mass Communication (86th, Kansas City, Missouri, July 30-August 2, 2003). Communication Theory & Methodology Division.
PUB DATE	2003-07-00
NOTE	429p.; For other sections of these proceedings, see CS 512 480-498.
PUB TYPE	Collected Works - Proceedings (021) -- Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC18 Plus Postage.
DESCRIPTORS	Adolescents; Advertising; Athletics; Community; Higher Education; Interaction; Internet; Interpersonal Communication; *Journalism Education; *Journalism Research; Memory; News Media; *Research Methodology; Self Esteem; Sexually Transmitted Diseases; Television Research; Values; World Wide Web
IDENTIFIERS	Network Based Approach; New York Times; Organization Theory (Psychology); Political Communication; *Third Person Effect; Unidimensionality (Tests)

ABSTRACT

The Communication Theory & Methodology Division of the proceedings contains the following 14 papers: "Interaction As a Unit of Analysis for Interactive Media Research: A Conceptualization" (Joo-Hyun Lee and Hairong Li); "Towards a Network Approach of Human Action: Theoretical Concepts and Empirical Observations in Media Organizations" (Thorsten Quandt); "Community & Civic Values, Communication, and Social Capital: 'Bowling Alone' as a Product of Values and Communication" (Leo W. Jeffres, Jae-won Lee, Kimberly Neuendorf, and David Atkin); "None of the Above: Creating Mass Deliberation Without Discussion" (Ray Pingree); "An Amplification of Sensationalism: Comparing the Tonal Values of the 'New York Times' to the 'New York Post' Using Whissell's Dictionary of Affect in Language" (Paul Crandon and John J. Lombardi); "Do Mass Communication Studies Test Measures for Unidimensionality?" (John D. Richardson and Frederick Fico); "Is Herpes Entertaining?: An Application of Entertainment-Education to Text Information Processing Concerning STDs Among Adolescents" (Donna Rouner and Ralf Kracke-Berndorff); "Attributions of Advertising Influence and Negative Stereotypes Among First- and Third-Person Perceptions" (Don Umphrey and Tom Robinson); "Modeling Micro and Macro: A Multilevel Model to Predict Memory for Television Content" (Brian G. Southwell); "Assessing Co-Termination Between Coders in Unitizing Textual Data: A Multi-Response Randomized Blocks Permutation Approach" (Li Cai); "How General Principles of Organization Theory Explain Gatekeeping Decisions About News: A Revised View of the Field" (Hugh J. Martin); "Democratic Realism, Neoconservativism, and the Normative Underpinnings of Political Communication Research" (Erik P. Bucy and Paul D'Angelo); "The World Wide Web of Sports: A Path Model Examining How Online Gratifications and Reliance Predict Credibility of Online Sports Information" (Thomas J. Johnson and Barbara K. Kaye); and "Self-Esteem, Self-Affirmation and Threats to Self-Worth: Testing a Motivational Explanation for the Third-Person Effect" (Patrick C. Meirick). (RS)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 481 263

Proceedings of the
Annual Meeting of the Association for Education in Journalism and Mass Communication

86th, Kansas City, MO
July 30-August 2, 2003

Communication Theory & Methodology Division

- Interaction as a Unit of Analysis for Interactive Media Research: A Conceptualization...Joo-Hyun Lee & Hairong Li
- Towards a Network Approach of Human Action: Theoretical Concepts and Empirical Observations in Media Organizations...Thorsten Quandt
- Community & Civic Values, Communication, and Social Capital, "Bowling Alone" as a Product of Values and Communication...Leo W. Jeffres, Jae-won Lee, Kimberly Neuendorf, David Atkin
- None of the Above: Creating Mass Deliberation Without Discussion...Ray Pingree
- An Amplification of Sensationalism...Paul Crandon, John J. Lombardi
- Do Mass Communication Studies Test Measures for Unidemensionality?...John D. Richardson, Frederick Fico
- Is Herpes Entertaining? An Application of Entertainment-Education to Text Information Processing Concerning STDs among Adolescents...Donna Rouner, Ralf Kracke-Berndorff
- Attributions of Advertising Influence and Negative Stereotypes Among First-and Third-Person Perceptions...Don Umphrey, Tom Robinson
- Modeling Micro and Macro: A Multilevel Model to Predict Memory for Television Content...Brian G. Southwell
- Assessing Co-termination Between Coders in Unitizing Textual Data: A Multi-response Randomized Blocks Permutation Approach...Li Cai
- How General Principles of Organization Theory Explain Gatekeeping Decisions About News: A Revised View of the Field...Hugh J. Martin
- Democratic Realism, Neoconservativism, and the Normative Underpinnings of Political Communication Research...Erik P. Bucy, Paul D'Angelo
- The World Wide Web of Sports: A Path Model Examining How Online Gratifications and Reliance Predict Credibility of Online Sports Information...Thomas J. Johnson, Barbara K. Kaye
- Self-esteem, Self-affirmation and Threats to Self-worth: Testing a Motivational Explanation for the Third-person Effect...Patrick C. Meirick

BEST COPY AVAILABLE

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

J. McGill

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 512 488

Running Head: INTERACTION AS A UNIT OF ANALYSIS

INTERACTION AS A UNIT OF ANALYSIS FOR INTERACTIVE MEDIA RESEARCH: A CONCEPTUALIZATION

Joo-Hyun Lee & Hairong Li

leejooh2@msu.edu / hairong@msu.edu

Michigan State University
Department of Advertising, CAS 309
East Lansing, Michigan 48824
phone. 517-214-1051

Prepared for the 2003 Annual Convention of the Association for Education in Journalism and Mass Communication
Kansas City, MO
July 30 – August 2, 2003
Communication Theory and Methodology Division

Keyword: interactivity, interaction

Joo-Hyun Lee is doctoral candidate of the Department of Advertising, Michigan State University.
Hairong Li is associate professor of the Department of Advertising, Michigan State University.

Interaction as a Unit of Analysis for Interactive Media Research: A Conceptualization

Abstract

This conceptual paper proposes interaction as a unit of analysis in interactive media research. Ambiguity of interactivity as a core concept has been identified. With a delineation of the similarities and differences among interactivity, reaction, and interaction, this paper presents a new definition of interaction. The superiority of the interaction concept over interactivity is explained, along with the antecedents and consequences. Research propositions and hypotheses are proposed for use of interaction in future interactive media research.

Interaction as a Unit of Analysis for Interactive Media Research: A Conceptualization

Academic discussions on interactivity have a short history. Although the construct "interactivity" has not always been clearly labeled as such, the idea existed in traditional communication models that emphasized feedback processes in many communication studies, and it was after the advent of many new media that interactivity became a widely popular topic for researchers. In particular, the emergence of the Internet and the development of relevant technologies in the late 1990s have brought about a variety of interactivity studies as the Internet was almost the first medium in history that featured a full range of interactivity with multimedia content.

Currently, two popular paths of interactivity definitions exist in communication studies. One path focuses on interactive features (e.g., Massey & Levy 1999; Newhagen & Rafaeli 1996; Rice 1984; Rogers 1986), whereas the other puts more emphasis on the individual audience's perception on the interactivity (e.g., McMillan 2000b; Newhagen 1998; Wu 1999). When focusing on the feature-oriented perspectives, it is obvious that the level of interactivity will keep increasing in various media with the rapid deployment of new technologies. It might pose a problem for interactivity studies because elements or features that were once regarded as very interactive could lose their innovativeness although they still might be interactive by definition. Consequently, it would be risky to adhere to certain interactive features in examining the effects of interactivity. Interactivity perception, or perceived interactivity that represents the other path of interactivity definitions would be less affected by improving technologies because it emphasizes on the audience's perspectives rather than the (interactive) features themselves. But taking this path to examine interactivity has its own problems as well.

This paper recognizes how interactivity construct might be vulnerable to various elements surrounding the construct (e.g., technological development, audience characteristics), and how it might be limited in explaining different conditions of media and audiences. This paper is an effort to explain those limitations and propose an "interaction" as a complementary construct that could better explain audience responses from various media technologies. The following sections will address definitions of interactivity and discuss the weaknesses.

Interactivity Definitions: Different Approaches

Although a number of studies have tried to define and explain interactivity from different angles, some suggest that the construct still needs clearer conceptualization (Brody 1990; Heeter 1989; Morris and Ogan 1996; Pavlik 1996; Rafaeli 1988). As Chen (1984) indicated, looking "beyond the technology of each new medium…will enable theoretical progress that does not stop at the borders of each new medium" (p.284). Likewise, the core concept of interactivity (along with its roles) would not change when the definition is clear and embraces different media.

Interactivity is generally regarded as a characteristic of a communication system (Williams, Rice, and Rogers 1988), a communication process (Chen 1984; Rafaeli 1988), or a combination of both (Heeter 1986, 1989). Kiousis (2002) summarized various descriptions of interactivity and accordingly proposed a third category of definitions (i.e., communication setting), but still most definitions of interactivity can be divided into two groups – feature-oriented definitions and perception-oriented definitions. In the beginning, researchers focused on the interactive features of a medium and presented feature-oriented definitions (e.g., Heeter 1989; Rice 1984), but later studies started to recognize the effect of an audience's perception of interactivity (e.g., McMillan 2000b; Newhagen 1998; Newhagen, Cordes, and Levy 1995; Wu 1999).

Feature-oriented perspectives define interactivity as the capability of new communication systems to *talk back* to the user (Rogers 1986). Similarly, it was described as the extent to which communication reflects back on itself, feeds on, and responds to the past (Newhagen and Rafaeli 1996). Another important condition for the interactivity is that communication roles between sender and receiver must be interchangeable (Williams et al. 1988). Synchronicity is another characteristic of interactivity, but there is a general consensus that synchronicity alone does not make a necessary nor a sufficient condition for interactivity to occur (Fortin 1997; Heeter 2000).

Heeter (1989) employed the concept of control in describing interactivity, and pointed out that users in interactive condition would have more control over the information to which they wish to be exposed. Using Heeter's (1989) multidimensional definition of interactivity, Massey and Levy (1999) examined the level of interactivity in a web site based on the presence of interactive features (e.g., e-mail links, chat rooms). Although there are slight differences in the definitions of feature-oriented interactivity, it should be noted that most of them emphasized exchange and mutuality, as Rice (1984) would note, "new media… facilitate interactivity among users or between users and information" (p.35).

Newhagen, Cordes, and Levy (1995) was one of the first studies that highlighted the psychological dimension of interactivity, which emphasized the communication participant's perception of his/her interactivity and of the receiver's interactivity. Steuer (1992) emphasized the individual's experiential aspect of interactivity, and defined interactivity as the extent to which users can participate in modifying the form and content. Newhagen (1998) argued that although the medium's features may be important in facilitating interactivity, the way that individuals use a medium would explain the interactive process better.

In short, the perception-oriented approach emphasizes the possible differences in the level of interactivity perceived by different audiences of a same medium. McMillan (2000b) documented that interactivity resided largely in the user's perception. In particular, McMillan (2000b) examined whether the interactive features would influence users' perception of interactivity, and found a very weak relationship. She also found that both interactivity features and perceived interactivity had a positive influence on users' attitude toward a web site (McMillan 2000b), yet the perception was a stronger indicator than the features (McMillan 2000a, 2000b).

The distinction between feature- and perception-oriented perspectives is important as they may generate different outcomes, and that interactivity may vary within the same medium for different vehicles and for different users within a same vehicle. For example, the web is regarded as a highly interactive medium but some web sites do not offer as much interactivity as others. Television is regarded as less interactive, but some audiences participate in interaction with programs or vehicles that offer such chances as call-in discussions, ARS (Audience Response System) quiz shows, and so on. Similarly, Rafaeli (1990) noted that traditional mass media audiences are more actively engaging in the communication with the media by writing letters to the editor and calling into on-the-air talk shows.

In an attempt to reconcile the different approaches of feature- and perception-oriented definitions of interactivity, Kiousis (2002) proposed a new conceptual definition, which reads, "the degree to which a communication technology can create a mediated environment in which participants can communicate (one-to-one, one-to-many, & many-to-many)…" "…it additionally refers to the ability of users to perceive the experience to be a simulation of interpersonal communication and increase their awareness of telepresence (p. 281"). Although this definition encompasses both approaches, it remains unchanged that feature-oriented descriptions cannot explain the individual audience's different perception of the same medium. Following this approach, televisions are always less interactive than e-mails. Perception-oriented interactivity definitions would help describe the differences in individual audiences, but it would not be very helpful in categorizing the various media based on the degree of interactivity. Consequently, it would be less useful to examine the level of interactivity in communication media. Despite these shortcomings, both feature-oriented and perception-oriented perspectives provide some helpful insights in understanding the interactivity construct and in establishing the interaction construct.

Interaction and Interactivity

Regardless of different definitions and conceptualizations of interactivity, the interactivity construct centers on the basic notion of human actions, reactions, or interactions. Therefore, examining interactivity without taking an individual's interaction into account would be far less useful because comparing audiences who interacted with a medium with those who did not could yield different results in terms of their response outcomes such as attention, comprehension, and the level of involvement.

Several reasons can be presented to explain the importance of interaction. First, a communication medium presents different conditions for interaction where different amount of interactivity might be selected by its audience. (e.g., *Jane and John both wanted to call the radio station to participate in a quiz, but only Jane could, because John had some other things to do.*) Second, the amount of perceived interactivity might vary for different audiences using a medium.

(e.g., *John only knows how to send and receive e-mails, while Jane is running an online virtual community.*) Finally, different audiences may have different levels of tendency to interact with the medium. (e.g., *John would never buy anything from the Home Shopping Channel, but Jane would buy anything that seems reasonably priced.*) In a similar vein, Heeter (1989) noted that different media systems require different levels of user activity. She pointed out that although users are always active with media to some extent, some users are more active than others and some media are more interactive than others. For example, e-mails are regarded as highly interactive medium, but some people may use its interactive features less than other people do. Conventional television and radio are regarded as non-interactive mass media. However, in some cases, audiences can enjoy interactivity by participating in live discussions.

Feature-oriented definitions of interactivity describe a medium's capacity to facilitate audiences' interactions. Perception-oriented definitions of interactivity illustrate an individual's perception of the responsiveness and control engendered by interactive features. Emphasizing the importance (and differences) of interaction would not deny that there are differences in the level of interactivity across different media, or that different individuals may experience varying levels of interactivity. Instead, it points out that there is a difference between "having a chance to interact" and "actually participating in the interactive communication process (i.e., interacting)" Feature-oriented definitions would describe a certain situation or a medium to be interactive when it provided chances (or options) to interact to its audiences. Perception-oriented approaches may presume that everyone in a certain situation perceived a similar level of interactivity, failing to distinguish those who interacted from those who did not.

The difference can be explained only by examining people's actual interactions. However, no studies to the researcher's knowledge have tried to distinguish interaction from interactivity, and few examined the role of interaction in audiences' information processing. Before proceeding to interaction conceptualization, another classification of interactivity—*person interactivity* and *machine interactivity*—needs to be reviewed as it will help understanding the conceptualization of interaction.

Person Interactivity and Machine Interactivity

Steuer (1992, p.84) explained machine interactivity as "the extent to which users can participate in modifying the form and content of a mediated environment." Steuer (1992) also emphasized the role of media (in a model of mediated communication) as a facilitator of person-to-person interaction by noting that media serve as a conduit in which message senders and receivers could interact. Hoffman and Novak (1996) viewed interactivity in terms of feedback, and explained that a computer-mediated environment enables users to *communicate through* the medium (i.e., person interactivity) and to provide or interactively *access* media content (i.e., machine interactivity). In other words, Hoffman and Novak (1996) stated that interactivity could be *through* the medium (emphasizing the human communication process mediated by machine – person interactivity) or *with* the medium (emphasizing the human interactions with the content – machine interactivity). The interactivity features of the medium are central in machine interactivity, since they directly *enable* the interactions. The machine would play the role of a communicator.

On the other hand, the features would be less important in person interactivity, because in this condition the machine only *facilitates* human interactions. The machine performs only as a mediator. Excluding the unmediated interpersonal communication, which is not the focus of this study, it can be said that a machine or a medium always plays a certain role – a communicator or a mediator. It should be noted that the above discussions on person and machine interactivity do not describe either interactive features or an audience's perceptions of interactivity. Although not clearly stated, the above discussions focus on the aspects of actual interactions occurring among users (i.e., person interactivity) or between users and media (i.e., machine interactivity).

Feature-oriented interactivity characterizes a medium's role in generating interactivity. Perception-oriented interactivity focuses on human's feeling. On the other hand, interaction refers to a behavior-oriented *communication process* no matter it is between people or between people and media. This paper examines the role of interaction, and discusses the degree of a medium's interactivity based on its potential to generate human interactions. In order to do so, a more detailed conceptualization of the interaction concept is needed.

Conceptualizing Interaction

Heeter (2000) conceptualized interactivity while taking interaction into account. Her approach provided a valuable starting point for conceptualizing interaction in this paper. Primarily, Heeter (2000) suggested that the concept of "interaction" would encompass a wide range of internal responses of an audience, which include thinking, feeling, attention, interpretation, and intention. In the beginning, she included every human action with an object as an interaction. This means that web users' simple mouse movements, data inquiry, and their cognitive/affective responses be interpreted as interactions. Then, she limited the interactions to the actions "physically observable" to separate the concept from such internal processes as perception, motivation, emotions, and so on. It was noted that those internal dimensions of interaction were "not subject to direct observation," and Heeter (2000) drew a line between interaction and other (internal) responses, and defined interaction as "an episode or series of episodes of physical actions and reactions of an embodied human with the world, including the environment and objects and beings in the world (Heeter 2000)."

This paper proposes to further refine Heeter's (2000) interaction definition. According to Heeter's (2000) definition, television audiences' flipping channels can be understood as an interaction. Also, a magazine reader's particular reading habit can be interpreted as an interaction as it is observable. But these types of interactions have a limited capacity to explain a medium's interactivity (or interaction-generating potential), although they are related to personal characteristics and tendency to interact. Therefore, it would be helpful to find a way to systematically differentiate these types of interactions from other types of interactions such as communicating with the message sender or interacting with a medium's contents.

Simply speaking, channel flipping actions and particular reading habits can be said to reflect how an audience consumes, processes, and *reacts* to the stimuli provided by the medium. Such activities are not sufficient to be called as an interaction. Rather, they are closer to reactions. Thus, the attempt of the current study to refine Heeter's (2000) interaction concept starts from distinguishing interactions from reactions.

The Merriam-Webster dictionary describes interaction as a "*mutual* or *reciprocal* action or influence" or "to act *upon one another*." Reaction is defined as "the act or process or an instance of reacting (which is "to *respond to a stimulus*"); a response to some treatment, situation, or stimulus, and; *bodily response* to or activity aroused by a stimulus." Interestingly, the heart of Heeter's (2000) interaction conceptualization – the observable nature – is found under the description of reaction. A clue for differentiating interaction from reaction can be sought from many interactivity definitions, which emphasize *mutuality* and the aspect of *two-way communication*. For example, Rafaeli's (1988) definition of interactivity is based on the "responsiveness" of a communication counterpart in the communication process. He noted that for a communication to be fully interactive, the sender-receiver roles must be *interchangeable.*

Based on this, this paper proposes a refined conceptualization of interaction using the concept of interchangeability, and it is *observable physical actions an audience performs in response to messages (content) provided through a medium which alter the content being provided and/or which communicate with the sender (publisher), either synchronously or asynchronously.*

The mutuality in the context of communication exchanges was established with the new conceptualization, but there is still one more issue that calls for further investigation. As mentioned earlier, certain interactions (e.g., channel flipping) are different from other interactions (e.g., writing back to a magazine), and the new definition by itself falls short in fully explaining the difference, as it includes both the publisher and the medium for the communication counterpart to which the feedbacks can be sent.

The answer may be found from the aforementioned rationales of person interactivity and machine interactivity. Talking back to a publisher or sending information in a web site may be understood as a function of person interactivity as the audience's interaction would reach the original message sender. The communication counterpart for this kind of interactions would be a person or an organization. This type of interaction embodies higher interchangeability, and can be labeled as the *human interaction* (with person or organization). On the contrary, such interactions as channel flipping, reading habits, recording a program, or increasing the volume represent the interactions that hardly ever reach the sender, and they can be understood as a function of machine interactivity. Usually, it involves no human communication counterpart, and the world is oblivious to this interaction made by a user. This type of interaction illustrates interactions with the medium or content, and can be labeled as the ***content interaction***.

Both types of interactions feature the core aspect of the interaction definition provided in this paper (i.e., roles interchangeability). The only difference between these types of interactions is in the communication counterpart – (medium-mediated) person versus the medium itself. Table 1 illustrates the difference between human- and content-interaction.

Interaction – Unit of Analysis

Interaction was defined as having physical observability and interchangeability of the sender-receiver roles. For a concept to be used in comparison with other constructs, it should be measurable, but the units of analysis for interaction may take different forms in different media. For television and radio audiences, writing a letter to the station can be an example of interaction. For web site visitors, clicking toward or away from certain web elements could be regarded as an interaction. Despite the differences, the behavioral patterns and accompanied goals would be similar across different media.

There are two ways to classify these patterns. First, interactions could be classified based on the time of media consumption – live consumption of content (e.g., click/volume increase), delayed consumption (e.g., save/record), and avoided consumption of content (e.g., closing a browser window). Second, interactions could be classified using an acceptance – avoidance dimension. Interactions of complete acceptance would include clicking *into* the web pages, saving the content for later use, bookmarking a web page, and increasing volume of the television set. On the other hand, interactions of complete avoidance would include closing a web browser window, clicking *away* from a web site, ignoring an e-mail, changing a channel, and turning off the equipment.

This paper treats not only acceptance but also avoidance as an interaction – only in an opposite direction and intention. However, some avoiding interactions might be interpreted as reactions as well. Note that the interaction definition in this paper described that the primary difference between interaction and reaction lies in the existence of interchangeability. Here is an example. There is a person viewing a web page and he sees a pop-up ad. When we focus on the person's web page consumption activity, his avoidance of the pop-up ad may be interpreted as a reaction to an interfering stimulus because the ad was interrupting his content consumption. In this case, the purpose of his avoidance or reaction was to continue with the content consumption. However, when examined from an advertiser's perspective, the same activity may be interpreted as an interaction. The advertiser places a pop-up ad to attract consumer's attention. When the ad is immediately closed, it represents the user's interaction because the action sends a message of avoidance to the advertiser. In this case, the same avoidance is interpreted as an interaction "away from" the stimulus.

Deciding between reaction and interaction in this example would depend on identifying the existence of interchangeability. The avoidance described above can be called a reaction when we focus on the person's content consumption, because the person does not send any messages to the content provider (i.e., content provider of the web page he was viewing). On the contrary, the same avoidance can be called an interaction when we focus on the person's response to the ad (and the user sending out the avoidance message to the advertiser). In this case, the content of the web page is not of concern.

Because the web is computer-based and generally considered to be more interactive than other mass media, interactions in the online environment would have a unique characteristic – ease of measurement. For instance, interactions on the Internet can be represented by audiences' clicking actions. Chatterjee, Hoffman, and Novak (1998) used visit duration and the number of pages visited as a possible measure of consumer interaction with web sites and banner ads. But the visit duration in web sites might be problematic when used alone. First, audiences' time spent in viewing web sites overlaps the number of pages viewed. Second, this measure can easily suffer from confounding variables such as the speed of connection, individual differences in comprehension rate, and the particular situation in which the person is browsing the web sites (e.g., concentrating on the content vs. doing something else at the same time). Although visit duration might be suitable for some experimental studies conducted in a computer lab, it would not be an appropriate measure of interaction in most cases.

Therefore, this paper recommends (1) the number of web pages visited by an audience member, (2) the number of clicks made to the overall hyperlinks (including ads) in a single web page or within a whole web site, and (3) the number of clicks made to a particular hyperlink as more valid measures of online interactions. Figure 1 illustrates the dimensions of interaction measures proposed in this paper.

Significance of Interaction

Stewart and Ward (1994) recommended that advertising studies should change the focus from analyzing media stimuli (and their impact) to exploring the way audiences interact with the media. The new definition presented in this paper will provide a means to more closely associate the concept of interaction with that of interactivity. Also, it will allow us to use the interaction concept as a means to examine the relationship between media interactivity and advertising effectiveness. For example, it has been believed that a mere exposure of an advertisement to the consumer is one of the key objectives of the advertisers. Current industry practices echo this notion as they employ pop-up ads, and as their online advertising pricing policies are based on reach and frequency. However, it should be noted that the fundamental goal of advertisers is to take audiences to the advertiser's web site or to generate sales. In other words, the more important goal of advertisers is to generate consumers' positive *interactions* with the ads.

Interaction represents the audience's goal as well. That is, a consumer's (series of) interactions in media use may be interpreted as efforts to achieve his or her goal in media consumption. The interactivity of a medium must be designed in a way that can help audiences achieve their goal in order to generate positive interactions, and therefore, content providers should always consider the reasons for a user's medium consumption.

Antecedents and Consequences of Interaction

Antecedents

Examining the effect of interactivity perceptions, McMillan (2000a) found that consumers' positive attitudes toward the web site (A_{ST}) better predicted their subsequent actions than interactivity perceptions do. Considering that McMillan (2000a) and Wu (1999) found that consumers' perceived interactivity affected their A_{ST}, it means that consumers' perception on interactivity has an indirect effect on their subsequent actions via their attitude. McMillan (2000b) further explained that the direct influence of interactivity perception on consumers' future actions was only partial and mostly limited.

However, the conceptual difference between *consumers' actions* in McMillan (2000a, 2000b) and the *interaction* concept developed in this paper should be noted. McMillan's (2000a, 2000b) *actions* referred to those that were limited to favorable reactions to the web site (e.g., telling about the web site and purchasing from the site), whereas the *interaction* in this paper is rather neutral in its nature. McMillan's (2000a, 2000b)'s actions could be rather easily predicted by attitude because both attitude and actions were measured based on consumers' favorability and only a few types of favorable actions were examined. Therefore, a direct application of McMillan's (2000a, 2000b) rationale (i.e., perceived interactivity $\rightarrow A_{ST} \rightarrow$ actions) might be problematic as it does not include the negative (inter)actions. Consumers do not have to be favorable to the web site in order to interact. For instance, they might not like the web site (e.g., online store with a poor interface), but they will still interact (e.g., purchase a product or browse for further information) when they find a reason to interact (e.g., cheap price).

In short, the interaction will not occur only because someone likes the content of the medium. Rather, interaction will take place when the user sees a certain benefit in making the interaction. Other possible reasons that would make it difficult to use the attitude as a precursor of interaction is the fact that interactions occur on specific elements in a medium (e.g., online chat rooms, ads, contents in need, etc.). Each element can affect the overall level of the audience's attitude toward the we site based on the audience's purpose of the web browsing, but the overall attitude toward the we site will not be able to perfectly explain whether the audience would interact with a specific element.

Considering that the interactivity construct originates from the basic principle of interaction, it can be thought that the interactivity perception and features will predict audiences' interaction with the web site better than the attitude. However, it should be noted that the perception and features of interactivity in media would only increase the *likelihood* of interaction, and that the perception and features would not *cause* interactions. People's perception or the feature of interactivity is like a well. If there is a well, people will come and drink from it. But it is hard to say that the well itself is the reason for the drinking behavior. Its presence might have increased the likelihood of drinking from that site, but few will drink water only because there is a well. In other words, the cause of the action cannot be represented by the presence of the well, but by the true driving motive of the action – thirst.

Similarly, interactivity features and perception may increase the chance of interaction, but they cannot represent the true cause. The true cause can be represented by people's needs. Novak, Hoffman, and Yung (2000) described that people's experience in web navigation can be categorized as ritualized experience and goal-directed experience. They explained the ritualized (or experiential) navigation as a less goal-oriented experience that is more oriented toward the experience itself. However, even when a user seems to be navigating the web without an explicit purpose, it can be understood that the user still has the implicit needs and goals such as killing time. Therefore, the media, contents, or the specific elements will generate more interactions when they better fulfill these needs and goals. On the other hand, the interactivity features and the perception would only *facilitate* interactions, and increase consumers' likelihood to interact, rather than representing the cause of interaction. From this, the following research proposition need be tested:

P1. The number of interactivity features in a medium will be positively related to the amount of interactions made by the user.

P2. The level of user's perceived interactivity in a medium will be positively related to the amount of interactions made by the user.

The above propositions aim to test the roles of features and perception in increasing the user's *likelihood* of interaction. That is, a direct comparison between high- and low-interactive feature conditions or between high- and low-perceived interactivity conditions might yield significant differences in the amount of interactions when the user's goal is controlled. Furthermore, it is possible that the users with different amount of interactions might have experienced the similar level of perceived interactivity. In this case, the differences in the amount of interactions will demonstrate the significance of the interaction measure.

Another factor that is expected to increase the chance of interaction is consumers' cognitive intensity in processing information. The more attention the consumer pays to the stimulus or the medium, the more likely the consumer is to show interactions. Supporting this argument, Cho (1999) found that an online consumer's interactions with banner ads (i.e., clicking) were related with the consumer's level of involvement with the advertised product. Furthermore, involvement was defined based on the level of attention and depth of processing (Gardner, Mitchell, and Russo 1978; Leigh and Menon 1987).

Some studies have employed the concept of flow to illustrate human-computer interactions (Csikszentmihalyi 1990; Trevino and Webster 1992; Webster, Trevino, and Ryan 1993). Novak, Hoffman, and Yung (2000) explained that consumers' flow experience during online navigation would make the consumers more involved in their navigation activity, which would let them more focus on their interactions. They conceptualized interaction as an "exploratory behavior," and showed that the flow experience had a significant influence on the exploratory behavior (Novak et al. 2000). Similarly, Berthon and Davies (1999) examined the effect of flow on people's interaction with the web site, and found a significant relationship.

Although the rationales regarding antecedents of interactions presented in those studies (Berthon and Davies 1999; Cho 1999; Novak et al. 2000) seem to be appropriate, the measures of interaction operationalization used in those studies are different from the conceptualization that this paper proposed. In particular, interactions were measured by people's intention to click (Cho 1999), intention to revisit the web site (Berthon and Davies 1999), and the general tendency to interact during online navigation (Novak et al. 2000). These measures do not correctly represent interaction because none of them are based on people's actual behavior. The intention or the tendency to click partially explains the likeliness to interact, but it must be noted that these intention-based measures have mostly been used in the context of consumer's intention to behave in a direction that is favorable to the advertiser/publisher. Considering the neutral nature of the interaction concept proposed in this study, the previous intention-based measures would not provide a perfect fit either for the chance of interaction or the interaction per se.

Cho and Leckenby (1998) explained online consumers' banner-clicking activity by investigating their underlying motivation. Although it was not empirically tested, they described that a consumer's clicking of banner ads could be explained by the motivations from advertising values (i.e., information/entertainment/usefulness), advertising characteristics (i.e., attention-/curiosity-generating), and user characteristics (consumer needs/involvement/learning motivation).

From the above discussions, the following hypothesis can be generated:

H1a. A consumer's level of involvement with the medium's content will be positively related to the amount of interactions made by the consumer with the medium.

H1b. A consumer's level of involvement with the product category featured in the ad will be positively related to the amount of interactions made by the consumer with the ad.

Consequences

Interactivity studies assume that a reciprocal and two-way communication is a commonly desired trait of media. Emphasizing exchange and mutuality, many interactivity definitions assume that the audience desires interacting with others (e.g., people, media, etc.). However, these assumptions are not shared by everyone. For example, Ha and James (1998) criticized the desired characteristic of interactivity to be unrealistic, and proposed that individual differences in communication needs should be considered. Lee and Lee (1995) also pointed out that interacting with a medium might be considered to be disturbing for certain audiences or contents. Neuman

(1991) argued that audiences might prefer not having to interact although having a choice of interactivity would be beneficial.

Interaction (or more specifically, *having to* interact) may be annoying to audiences in a certain situation, and it is due to the consequences interactions could cause. First, interactions require the person to pay more attention to the stimuli and the communication process as (s)he will practice an active control. As a result, the person will elaborate on the provided messages and experience deeper levels of cognitive processing. Cho and Leckenby (1998) argued that consumers' interaction with messages or advertisers was likely to generate active and intensive information processing. The intensified information processing can be interpreted as heightened level of consumers' situational involvement. That is, (a series of) interactions will produce self-generated thoughts because of the two-way nature of interaction. Participating in a two-way communication process means that the person exchanges messages actively rather than passively receiving them. These exchanges of messages and making decisions will intensify the communication process, which will heighten the level of their cognitive involvement occurring in the communication process.

Second, interactions represent the person's investments or efforts, and investments would stimulate attention and involvement. For example, when an audience picks out a favorite contestant while watching Fox's *American Idol* and votes for the contestant by calling the toll-free number, this interaction would make the audience pay more attention to the result and more involved with the program (e.g., greater desire for the contestant to win the match) compared to those who did not make such an interaction. Also, voting on an issue in a web site might generate similar effects and increase situational involvement with the web site. Consequently, continuing interactions would not only reflect the person's involvement levels of the object, but also reinforce the level of involvement unless the interactions are interrupted by other factors such as unsatisfactory results or distracting stimuli. Figure 2 illustrates the overall antecedents and consequences proposed in this paper.

From the discussions on consequences of interactions, following hypotheses can be generated:

H2. An audience's interaction with the ad or media content will increase the level of the audience's perceived interactivity.

H3. An audience's interaction with the ad will increase the level of the audience's situational involvement with the ad and the advertised brand.

H4. An audience's interaction with the ad will increase the level of the audience's recall of the advertised brand.

H5. An audience's interaction with the ad will increase the level of the audience's situational involvement with the medium in which the ad was placed.

Conclusions and Discussions

So far, this paper has discussed the concept of interaction, centering on its differences from interactivity features and perception. Interaction is an important and differentiated concept from interactivity. Focusing on the feature-oriented aspect of interactivity would be useful in explaining the differences of various media. Studies on perceived interactivity are noteworthy because they recognize the fact that the individual consumer's perception on interactivity might differ within a same medium. However, the feature-oriented interactivity cannot explain the difference among users. Perception-oriented interactivity cannot tell the difference of the users who interacted with the medium from those who did not. Furthermore, neither feature- nor perception-oriented interactivity can explain users' different levels of tendency to interact.

Interaction is another measure of individual differences. But as it represents an observable behavior, interaction will better explain the differences among different media. For instance, magazines are believed to be a less interactive medium than the Internet. Determining the interactivity levels of magazines and the Internet based on the number of interactive features is not an easy job because the features may not be directly comparable and because the audiences' perception of interactivity may differ. Comparing the amount of interactions generated by each medium may provide researchers with a better means to determine the level of interactivity among different media. Although the audiences of different media will demonstrate different forms of interactions, they may be categorized and analyzed according to the interaction typologies proposed in this paper (e.g., human- and content-interaction).

The interaction construct's observable nature will also better explain the audiences' differences in the level of perceived interactivity within a same medium. Perceived interactivity in a media environment is useful because it may predict the audience's responses better than interactive features (McMillan 2000a, 2000b). But perception may not accurately predict interaction occurrences because the audience may have different level of tendency and ability to interact. Another advantage of the interaction concept lies in the fact that interaction directly examines the principles that constitute both the feature- and perception-oriented interactivity descriptions, which are controllability and mutuality.

The goal of this paper has been to present a more refined conceptualization of interaction to aid in investigating its antecedents and consequences, and to motivate further research in this area. studying an effort to do so, this paper added a new characteristic (i.e., interchangeability) to the interaction concept that was previously provided by Heeter (2000). Communication technologies will evolve and new interactive features will appear in media environment. Rather than relying on those features to examine the impact of interactivity, communication researchers should become involved in investigating how the audience's interactions in a new media environment would make a difference in their media consumption pattern. This paper provided an interaction typology based on person- and machine-interactivity rationales, and it would represents a means to categorize a variety of interactions that are occurring in different media environments.

References

Berthon, Deon Nel Raymond van Nickerk Jean-Paul and Tony Davies (1999), "Going with the Flow: Web Sites and Customer Involvement," *Internet Research*, 9(2), 109-116.

Brody, E. W. (1990), *Communication Tomorrow: New Audiences, New Technologies, New Media*, New York: Praeger Press.

Chatterjee, Patrali, Donna L. Hoffman, and Thomas P. Novak (1998), "Modeling the Clickstram: Implications for Web-based Advertising Efforts," Manuscript at *eLab*, Vanderbilt University, [Available] http://ecommerce.vanderbilt.edu/research/papers/pdf/manuscripts/Clickstream-pdf.pdf

Chen, Milton (1984), "Computers in the Lives of Our Children: Looking Back on a Generation of Television Research," in Ronald Rice et al., *The New Media: Communication, Research, and Technology*, Beverly Hills, CA: Sage, 269-286.

Cho, Chang-Hoan (1999), "How Advertising Works on the WWW: Modified Elaboration Likelihood Model," *Journal of Current Issues and Research in Advertising*, 21(1), 33-50.

__________, and John D. Leckenby (1998), "Copytesting of Advertising on the WWW: Clicking Motivation Profile," in *Proceedings of 1998 Conference of the American Academy of Advertising*, American Academy of Advertising, 26-36.

Csikszentmihalyi, Mihaly (1990), *Flow: The Psychology of Optimal Experience*, New York: Harper and Row. Basic Books.

Fortin, David R. (1997), *The Impact of Interactivity on advertising effectiveness in the new media*. Unpublished doctoral dissertation, University of Rhode Island.

Gardner, Meryl Paula, Andrew A. Mitchell, and J. Edward Russo (1978), "Chronometric Analysis: And Introduction and an Application to Low Involvement Perception of Advertisements," in *Advances in Consumer Research*, Vol. 5, H. Keith Hunt, ed., Ann Arbor, MI: Association for Consumer Research, 581-589.

Ha, Louisa and E. Lincoln James (1998), "Interactivity Reexamined: A Baseline Analysis of Early Business Web Sites," *Journal of Broadcasting & Electronic Media*, 42(Fall), 457-474.

Heeter, Carrie (1986), *Perspectives for the Development of Research on Media Systems*, Unpublished doctoral dissertation, Michigan State University, East Lansing, MI.

__________ (1989), "Implications of New Interactive Technologies for Conceptualizing Communication," in *Media Use in the Information Age: Emerging Patterns of Adoption and Consumer Use*, Jerry L. Salvaggio and Jennings Bryant, eds., NJ: Lawrence Erlbaum Associates, 217-235.

__________ (2000), "Interactivity in the Context of Designed Experiences," *Journal of Interactive Advertising*, 1(1), [http://jiad.org/vol1/no1/heeter/index.html].

Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing*, 60(July), 50-68.

Kiousis, Spiro (2002), "Interactivity – A Concept Explication," New Media and Society, 4(3), 255-283.

Lee, Barbara and Robert S. Lee (1995), "How and Why People Watch TV: Implications for the Future of Interactive Television," *Journal of Advertising Research*, 35(6), 9.

Leigh, James H. and Anil Menon (1987), "Audience Involvement Effects on the Information Processing of Umbrella Print Advertisements," *Journal of Advertising*, 16(3), 3-12.

Massey, Brian. L and Mark R. Levy (1999), "Interactivity, Online Journalism, and English-Language Web Newspapers in Asia," *Journalism & Mass Communication Quarterly*, 76, 138-151.

McMillan, Sally J. (2000a), "Interactivity is in the Eye of the Beholder: Function, Perception, Involvement, and Attitude toward the Web Site," in *Proceedings of the 2000 Conference of the American Academy of Advertising*, ed., Mary Alice Shaver, East Lansing, MI: Michigan State University, 71-78.

___________ (2000b), "What is Interactivity and What does it do?" Paper presented at the *Annual Conference of Association for Education in Journalism and Mass Communication*, Phoenix, AZ.

Merriam-Webster Collegiate Dictionary (2002), [http://www.m-w.com].

Morris, Morris and Christine Ogan (1996), "The Internet as Mass Medium," *Journal of Communication*, 46(1), 39-50

Neuman, W. R. (1991), *The Future of the Mass Audience.* Cambridge, MA: Cambridge University Press.

Newhagen, John E. (1998), "Hitting the Agenda Reset Button for the Internet: The Problem of Matching Research with Development," Paper presented at the *International Communication Association Conference*, Jerusalem, Israel (July).

___________, John W. Cordes, and Mark R. Levy (1995), "Nightly@nbc.com: Audience scope and the perception of interactivity in viewer mail on the Internet," *Journal of Communication*, 45(3), 164-175.

___________ and Sheizaf Rafaeli (1996), ""Why Communication Researchers Should Study the Internet: A Dialogue," *Journal of Communication*, 46(1), 4-13.

Novak, Thomas .P., Donna L. Hoffman, and Yiu-Fai Yung (2000), "Measuring the Customer Experience in Online Environments: A Structural Modeling Approach," *Marketing Science*, 19(1), 22-44.

Pavlik, John V. (1996) *New Media Technology*. Needham Heights, MA: Allyn and Bacon.

Rafaeli, Sheizaf (1988), "Interactivity From New Media to Communication," in *Advancing Communication Science: Merging Mass and Interpersonal Processes*, Hawkins, Robert P., John M. Wiemann and Suzanne Pingree, eds., Beverly Hills, CA: Sage, 110-134.

____________ (1990), "Interacting with Media: Para-Social Interaction and Real Interaction," in *Mediation, Information and Communication*, B. Ruben and L. Lievrouw eds., New Brunswick: Transaction Publishers, 125-181.

Rice, Ronald (1984), "New Media Technology: Growth and Integration," in *The New Media: Communication, Research, and Technology*, Rice, Ron and Associates, eds., Beverly Hills, CA: Sage, 33-54.

Rogers, Everett (1986), *Communication Technology: The New Media in Society*, New York: Free Press.

Steuer, J. (1992), "Defining Virtual Reality: Dimensions Determining Telepresence," *Journal of Communication*, 42(4), 73-93.

Stewart, David W. and Scott Ward (1994), "Media Effects on Advertising," in *Media Effects, Advances in Theory and Research*, Jennings Bryand and Dolf Zillman, eds., Hillsdale, NJ: Lawrence Erlbaum Associates.

Trevino, Linda Klebe and Jane Webster (1992), "Flow in Computer-Mediated Communication," *Communication Research*, 19(5), 539-573.

Webster, Jane, Linda Klebe Trevino, Lisa Ryan (1993), "The Dimensionality and Correlates of Flow in Human Computer Interactions," *Computers in Human Behavior*, 9(4), 411-426.

Williams, Frederick, Ronald E. Rice and Everett M. Rogers (1988), *Research Methods and the New Media*, New York: The Free Press.

Wu, Guohua (1999), "Perceived Interactivity and Attitude toward Web Sites," in *Proceedings of 1999 Conference of the American Academy of Advertising*, Marilyn S. Roberts, ed., Gainesville, FL: 254-262.

Appendix

Table 1. Human- and Content-Interactions

Interactions are observable physical actions an audience performs in response to messages (content) provided through a medium which alter the content being provided and/or which communicate with the sender (publisher), either synchronously or asynchronously.

	Human Interactions	Content Interactions
	Based on person interactivity Interactions reach the sender.	Based on machine interactivity Interactions hardly reach the sender.
Interchangeability	High	Low
Communication Counterpart	Person or organization	Medium or content (No human counterpart involved)
Examples	Talking back to a publisher, sending information in a web site, etc.	Channel flipping, recording a program, increasing the volume, etc.

Figure 1. Dimensions of Interaction Measurement

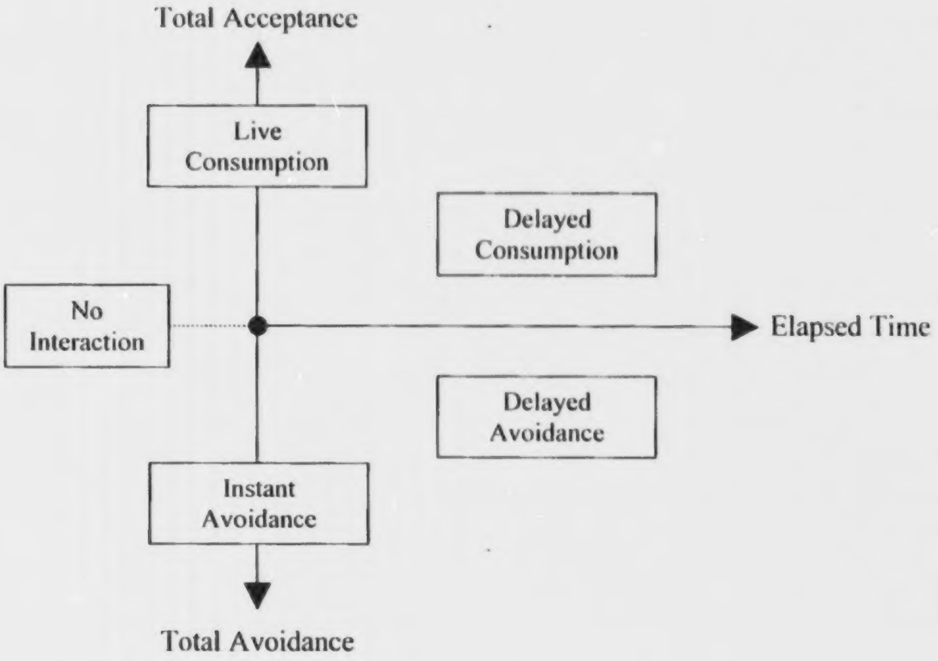
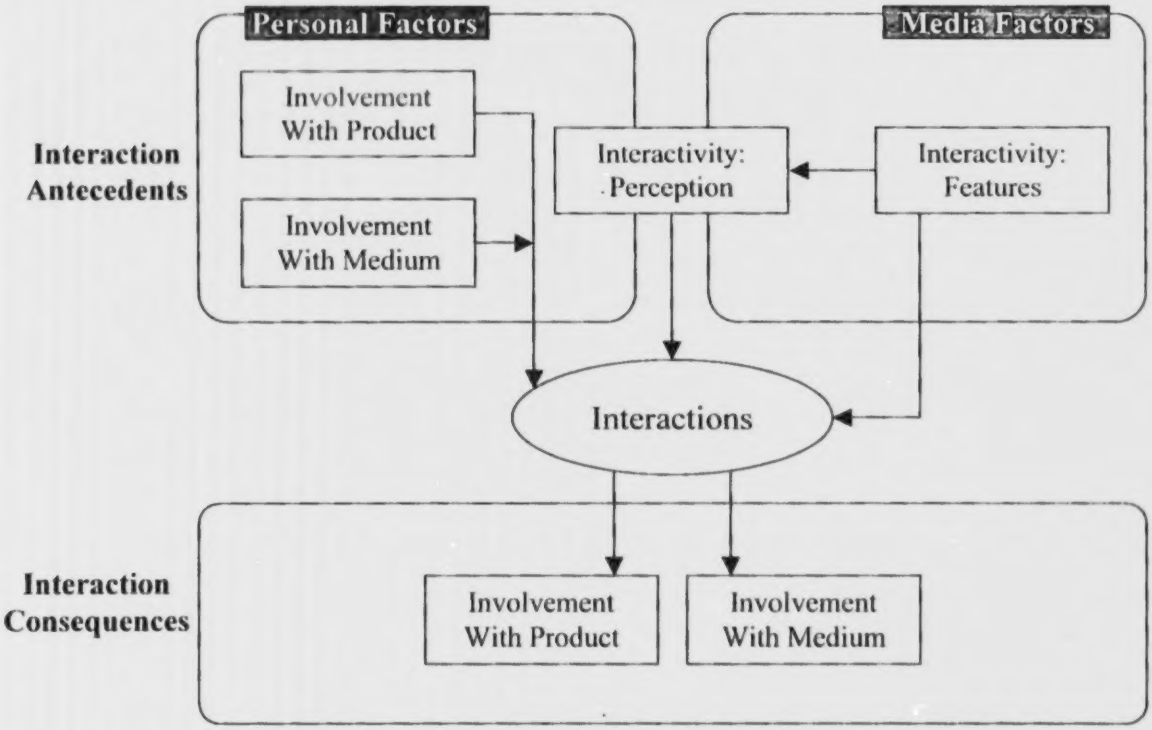

Total Acceptance
Live Consumption
Delayed Consumption
No Interaction
Elapsed Time
Delayed Avoidance
Instant Avoidance
Total Avoidance

Figure 2. Antecedents and Consequences of Interactions


Personal Factors
Media Factors
Interaction Antecedents
Involvement With Product
Involvement With Medium
Interactivity: Perception
Interactivity: Features
Interactions
Interaction Consequences
Involvement With Product
Involvement With Medium

BEST COPY AVAILABLE

Towards a Network Approach of Human Action

Theoretical concepts and empirical observations in media organizations

Contact:
Thorsten Quandt, M.A.
Institute of Media and Communication Studies
Technical University Ilmenau, Germany
P.O. Box 10 05 65
D-98684 Ilmenau
Fon: +49/3677/69-4669
Fax: +49/3677/69-4695
E-mail: thorsten.quandt@tu-ilmenau.de

Paper submitted to the
Communication Theory and Methodology Division

Towards a Network Approach of Human Action

Theoretical concepts and empirical observations in media organizations

Full Abstract

Although having a notable history in mathematics and the social sciences, it was not until recently that network approaches have become increasingly popular in the scientific community. Meanwhile, network approaches have successfully been applied to such diverse fields as the modeling of consumer behavior on the Internet, the search for DNA code or the uncovering of terrorist groups. Nevertheless, network approaches are quite uncommon in communication studies.

This paper argues that network concepts can be helpful in describing phenomena in the media. It therefore sketches the framework of a relational theory of human action and presents data from empirical observations in the news rooms of media organizations, which have been carried out on the basis of such an approach. During this 10-week project, we observed the behavior of six German online journalists and coded more than 11.000 of their actions. We found surprising similarities in the coded material which leads us to the conclusion that there are a number of associations and sequences in human action which can indeed be described and analyzed on the basis of network theory.

1. Introduction: Networks and social theory

Network approaches are attracting a lot of attention these days, and in particular from the general public. Just after September 11th 2001, the idea of networks has been widely discussed, primarily in reference to terrorist groups. Data mining algorithms based on networks algorithms have been applied in the search for Al-Qaida members. Similar mathematical models are used to identify consumer behavior on the Internet or patterns in the DNA code. On a more general level, network metaphors have been used to characterize modern society as a whole, even in newspaper articles and on TV. While many of these discussions are based on popular network ideas (and linked to similar phenomena like „the Internet"), some ideas actually stem from an academic debate that took place in the recent years.

There, one can identify several sources for such a discussion. The two major sources are:

(1) Mathematical concepts of networks derived from graph theory

(2) Sociological concepts based on the network metaphor

In the second case, the central term 'social connectivity' refers to a broad understanding of society being similar to a network – which arguably means: a network of interlinked agents (i.e. individuals or groups). Especially in media and cultural studies, some researchers focus on the role of media in connecting such agents.

While such an approach might be helpful in analyzing the relations between people and the media, it is not the only conceivable way to apply a network concept to human society (as explained in section 2). As an alternative way of employing network concepts, we want to present the idea of a network of action – a network that ultimately shapes the way we perceive and construct the world (section 3). We will argue that, on the basis of our individual actions, structures are emerging which can most likely be described in terms of a network of meaning. This theoretical concept is supported by data from an observational study of online journalism (section 4). There, it became quite evident that human actions may be characterized by a network of action elements, and also that suitable raw data taken during such observations can be analyzed by means of standard network analysis tools. In the last section, we will summarize the pros and cons of this new way of theoretical and empirical thinking suggested here (section 5).

2. Network approaches: Some roots

Network approaches are not as new as the current debate would lead us to believe. The concept of people forming a network is indeed an old one, and it was first introduced to sociology by researchers like Georg Simmel and Alfred R. Radcliffe Brown in the late 19th and early 20th century. They used these ideas to describe social phenomena and structures, but mainly on a metaphorical level. Empirical work, like the ethnographic studies of John A. Barnes on kinship and social structures pushed the sociological concept further ahead beyond its mere metaphorical meaning. Since then, the theoretical concept of networks in sociology and social sciences has been improved upon in many ways. In sociology as well as in economy, networks became a central concept for the description of structured phenomena: Williamson (1985) used this term to characterize a very efficient way of economic coordination, Perrow (1992) discussed the distribution of power and influence with the help of the network idea, Windeler (2001) applied the concept to organizations, and just lately Castells (2000) presented his vision of a network society, which has been discussed at lot since then, even outside the scientific community. These were just a few examples (cf. Scott, 2002, for a large overview of standard texts on networks).

On the other hand, there is another major field of network approaches which can be derived from the so-called "graph theory". The latter is the logical and mathematical basis for the formal description and analysis of networks and connections. A graph is a general type of structure which can be represented by elements (nodes) and its connections (links). The beginnings of graph theory date back to the late 18th century, starting with Leonhard Euler and his solution of the so called "Königsberg problem" (cf. Biggs, Lloyd, & Wilson, 1976). The mathematical graph theory was later refined in terms of a complex network theory, borrowing some ideas from chaos theory and the analysis of self-organizing systems (Barabási, 2002). With the increasing power of computer software, this kind of network analysis is becoming increasingly popular in many areas of research, ranging from the decoding of the human genome to the analysis of organizations or the uncovering of terrorist groups. The standard numerical tools include data mining packages and the application of artificial intelligence based analysis algorithms (cf. Klösgen & Zytkow, 2002).

It is not surprising that the applications of graph theory are manifold, due to the logical (and therefore empirically empty) quality of such a point of view. But it is quite of a surprise that the sociological point of view concerning networks is somewhat conservative when it comes to choose the types of phenomena that it should describe. Or to say it more clearly: the choice of network nodes. In most cases, the sociological network approach refers to society or groups as structures being similar to a web – forming a network of interlinked agents (individuals or groups). So the elements or nodes that appear in sociological network theories are human beings.

We would like to argue that this approach is far too narrow, and that network concepts can be applied to other social phenomena as well, especially to the basic category of human action. Actually, various companies on the Internet are already operating in the same direction. They try to model buying behavior using network algorithms: the nodes are the individual buying acts, which are connected to other buying acts, and in the end, there emerges a complex network of connected buying acts. This structure is what these companies are actually looking for in order to be able to predict consumer behavior. And there is already a general term for this kind of analysis: It is called "data mining" or "knowledge discovery" (cf. Klösgen & Zytkow, 2002).

In this paper we will argue in the same direction: we will not focus on a network of interlinked agents, neither on individuals or groups, but on networks of actions instead. Which does not mean that we want to leave out human beings, or that we want to suggest that networks of individuals or groups would not be a helpful concept. But we believe that choosing them as network nodes might not be the only promising way of applying network theory to social or cultural phenomena. In the following section, we would like to present some theory that will support our point of view.

3. Theoretical background: Towards a network approach of human action

When describing human actions, it cannot be enough to just label the types of individual acts that are being performed by a certain person (asking „what is s/he doing?“). There are a number of factors that determine the way in which these acts are finally embedded into the flow of action. For example there is the time and space framework („where and when does s/he do this?“), contact persons or relations among subjects („...in contact with which person?“), the material resources („...with the help of what type of resource?“) and the general sense making ‚location‘ of the act („...in which context?“). These elements may be looked upon as constitutional for human actions, and most of them have already been identified in the standard works on a sociological description of human action (cf. Schütz, 1981; Weber, 1972), and in more recent publications like Giddens' structuration theory (Giddens, 1997). While we can surely conceive of other elements as well, the elements described here may already be sufficient to characterize individual actions.

A figure illustrating these interconnected elements is shown below (fig. 1).

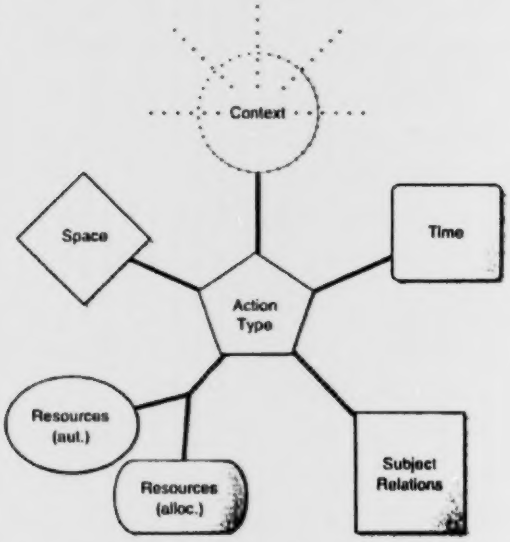


Fig. 1: An individual act as a star network of associated elements

In this figure, we clearly observe a network structure, but for good reason: The constituent elements of each act are linked by the action itself, and therefore they constitute a small star network. Without its central node, the network would cease to exist, whereas some of its outer elements might eventually be missing under certain circumstances (which is true for relations among subjects and resources that are not essential for each and every individual human action).

BEST COPY AVAILABLE

On the other hand, human actions do not exist as moments frozen in time. Instead, they are part of a constant flow in time, with one act followed by another. In our everyday life, we constantly do something, followed by another action, and based on a certain history of acting. Such a history of action is only possible because we perceive actions as being related to each other, and in particular when they take place in sequence. For example, we may assume that the majority of journalists know which steps to follow when they have to write an article: They know their sources for research, they remember possible starting points from earlier work on similar topics, they know when they have to talk to somebody, and they know when they should stop researching and begin with writing things and so on. They obviously remember single micro-steps as well as large coherent sequences. In common language we call that "experience". Sociologists and psychologist alike assume that humans remember actions through cognitive processes by means of what Schütz called a "stock of knowledge at hand" (Schütz, 2002, 153 f.). Which is a repertoire of basic rules being at our disposal in order to develop strategies for our future actions. In the language of the network theory, these rules are operating as connection rules, because they are able to describe the structures among various action elements like resources, types of actions, personal contacts, contextual information and the space and time framework itself. Therefore, this stock of knowledge is basically a huge network of relations which constitutes human memory and which lays the foundations for further human actions, thereby creating the very identity of a person performing those acts.

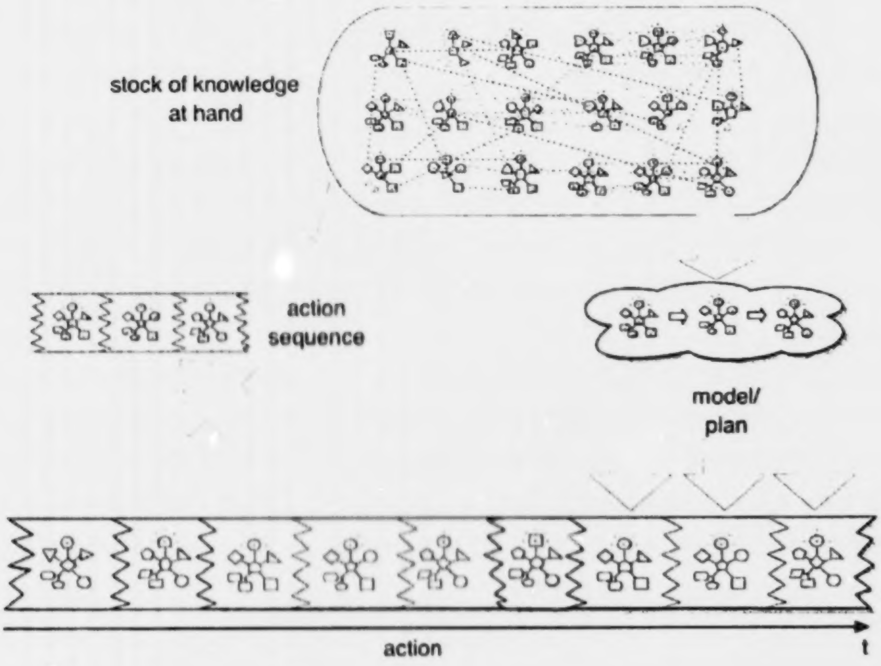


Fig. 2: Sequences of actions are being transferred to the stock of knowledge

BEST COPY AVAILABLE

Now what happens if several individuals, for example journalists in (different) newsrooms, have contact to similar subjects, similar resources, working under certain material conditions, and being confronted with similar actions? First of all, they will build up similar relations among certain action elements in their stock of knowledge. That does not necessarily mean that they are forming similar traces of memory in their brains. Actually this is highly unlikely, because the perceived actions usually relate to different elements in each individual stock of knowledge. But the important thing is that these subjects share the same relations. Let us take this paper as a simple example: As a reader you will perceive our words in one way or another. And the way in which you relate the information contained in this paper to your actual knowledge is a highly individual process, because we are all entering such a process with rather different memory structures. Nevertheless, you will share your relation to this paper with any other reader, even if s/he is thousands of miles away and lives in a totally different living environment.

So while the nature of links might vary, their relational qualities will basically be the same. They will also stay the same if people share parts of their stocks of knowledge through communication or through co-orientation.[1] There may not be direct contacts between all the initial action elements, but at least there remain some links. For example if you (as a reader) would tell a friend what is explained in this paper, say in a few days, this friend would most likely share a – somewhat weaker – tie to the present paper (cf. fig 3).

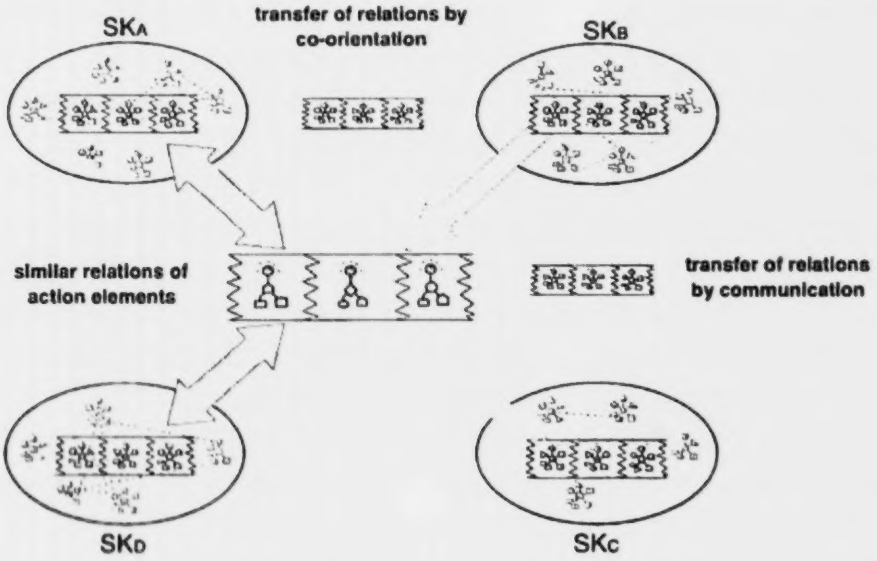

SKA
transfer of relations by co-orientation
SKB
similar relations of action elements
transfer of relations by communication
SKD
SKC

Fig. 3: Building networks of meaning through shared relations

---

[1] The similarity between those relations can be explained through co-orientation, i.e. orientation towards similar phenomena, and also through communication. However, similar structures are neither a necessary nor and an exclusive effect of communication. The latter can be described as a special type of action that transfers parts of 'ego's' memory structures into the stock of knowledge of 'the other'.

So through their everyday's practice, and through similar connections among actions, resources, contexts etc., people will build up comparable webs of sense-making relations (i.e. connections that allow for a complex behavior which creates options for subsequent human actions) – therefore, the individuals are actually sharing some 'meaning' (at least to a certain amount).

4. Empirical applications: Observational study of online journalists

4.1. Design of the study/Methodology

One of the advantages of network approaches is that they can easily be applied to empirical studies: After defining appropriate nodes and the relations among them, the structure of these networks may simply be described by graphs (which means that they form a logical structure which can be translated into a formal/mathematical language; cf. Scott, 2000, and Wasserman & Faust, 1994). The above-mentioned approach already provides us with the basic elements that may serve as categories for such empirical studies: The 'nodes' of individual acts can be operationalized for direct use in (observational) studies. Surely, such an approach may also serve as the basis for surveys, but observations seem to be most natural way to analyze human action/behavior.

Based on the theoretical approaches mentioned above, a large observational study could actually be realized. During a 10-week study in the newsrooms of 5 German online newspapers, the actions of 6 online journalists have been observed. The motivation for such a study was the idea that there might be some sort of professional rules evolving for this new area of journalism. At the time of the study, German researchers did do not know very much about 'real life' working conditions and the rules of online journalism (Neuberger, 2000, 37 f.). Therefore a closer look at the structures (the elements of actions and relations among them) shaping the everyday work of a journalist was certainly overdue.

The operationalization of the individual action elements (types of actions, context, space, time, resources, subject relations) resulted in a codebook containing about 250 numerical and symbolic codes that had to be memorized by the observer. During our observations, the flow of action was broken down into individual acts[2] and the acts themselves into the constituent elements which were itemized in the codebook (the graphics shown below should give us an impression of how this was done in principle, cf. fig. 4; different conditions/values of the individual elements are indicated by different shapes).

To provide a better understanding of the working environment, observational diaries were set up to write down open questions which could be answered during eleven interviews with the journalists and their editors in chief. In addition to that photographs of these workplace were taken and ground plans of the work places were drawn in order to get some impression of the working conditions of the journalists. But the core results were the coded observations. We obtained a data matrix with 11.671 acts (corresponding to 483 hours and 28 minutes of observation); each act consisted of about 50 variables that would describe its constituent elements in more detail.[3] Therefore the data basis for further analysis was huge.

---

[2] Acts were defined as being interconnected and coherent. Thus an act would end when at least one of its elements had changed. The question about the observed size of acts (the „granulation" of observation) is not answered by such a procedure – but this is was not the central question when we were looking for patterns, because relationships will be visible even when the size of observed acts does vary. The relational structure will actually stay the same (cf. Quandt, 2003).

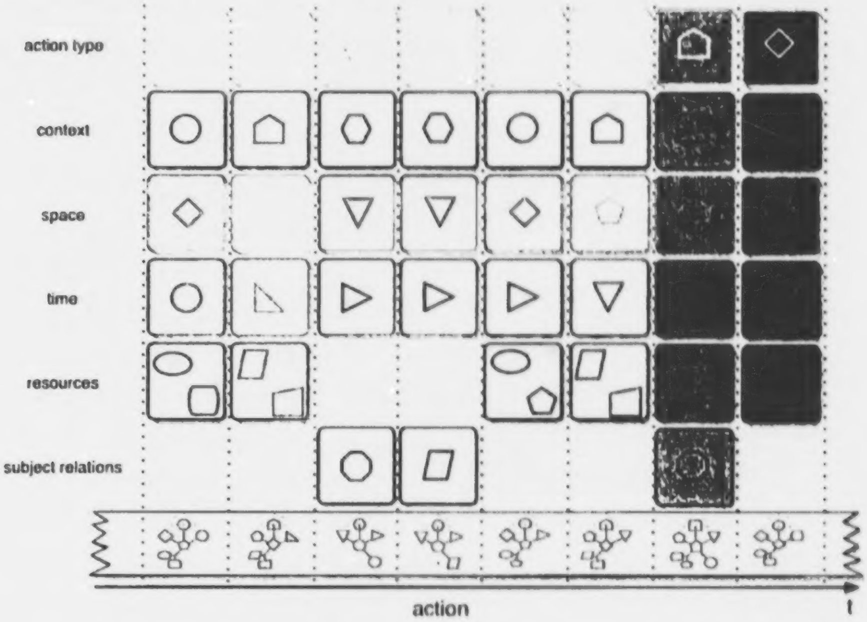


Fig. 4: Breaking down the flow of action into a data matrix

The aim of the study was to identify patterns among the relational data contained in this matrix. Or to put it in other terms: the aim was to find similarities and rules of action that might be typical for online journalists. Such patterns (work routines, rules of action) can evolve into different directions – first, there are frequent connections between different elements, which are called associations. Or one might find temporal patterns, which are called sequences. The main question concerning the latter type of connection was the following: Are there certain actions that follow other actions on a regular basis?

4.2 Results

4.2.1 Overall distribution of journalistic action

The results of this study reveal some striking similarities in the observed actions of the different journalists, although there were no direct contacts between the observed individuals – they all worked for different media organizations in different towns. Still the schedule of their working days, as well as their general rules of action and their use of resources (including technological devices) followed comparable patterns. There appear to be invisible ties between the individuals and their actions – which is "net-work" in both sense of such an expression. This is quite astonishing, given the fact that online journalism is a relatively new field, with no original tradition of its own that would make it different from journalism as a whole. Even the literature on this subject has not identified special rules of online journalism (although there really is a bulk of articles dealing with online journalism). The above mentioned idea of human actions as being shaped through the network of sense-making relations seems to be useful in explaining this fact: It is assumed that the similarity of relations leads to the formation of comparable structure building processes in the stock of knowledge, as well as among the observed actions.

---

[3] The high number of variables per case is partly due to individual coding for up to four resources, four (groups of) contact persons and several context variables.

Some empirical data from our study will give us an impression of the above-mentioned similarities. First of all, the overall distribution of types of actions was similar for almost all of the observed journalists (with the exception of one journalist who had a lot of technical tasks; this was actually due to the fact that he was the only online journalist in his media organization).

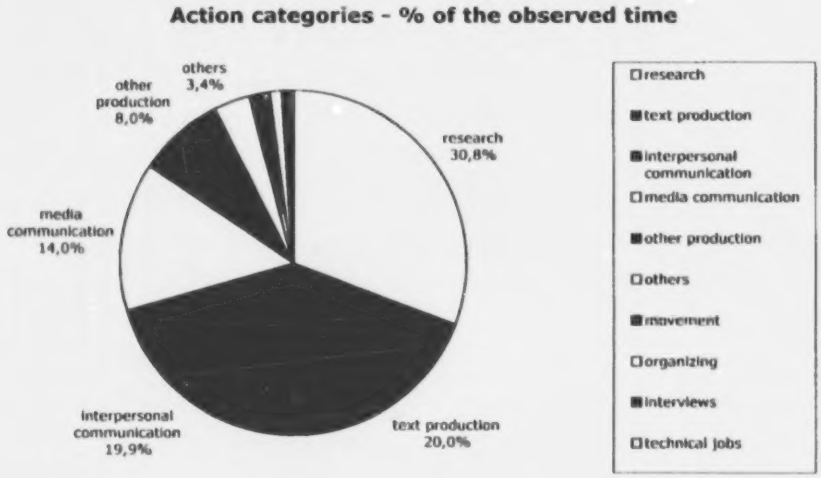


Fig. 5: Overall distribution of observed actions[4]

The above pie graph (cf. fig. 5) shows the overall distribution of time spent on different actions during a journalist's office hours. The biggest pieces are research, text production, interpersonal communication, the communication through media, and production jobs. That is roughly what one would expect from a journalist, although the high level of communication looks rather surprising (which mostly consists of co-ordination with colleagues, though – for example through organizational talks).

It is interesting to see the homogeneous distribution pies related to different journalists. The following viewgraph (cf. fig 6) compares the amount of time spent on the individual actions for two journalists of the Netzeitung in Berlin (NZ1 and 2), the Frankfurter Allgemeine Zeitung (FAZ) and the 'Tagesschau' in Hamburg (TS).

---

[4] Numbers for the smaller pieces of the pie have been omitted for the sake of clarity.

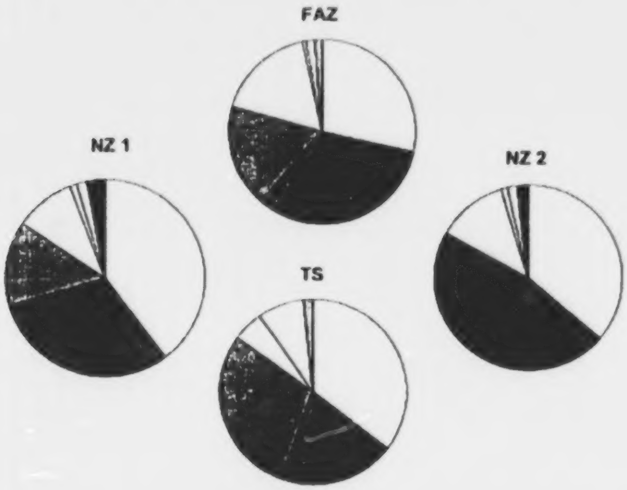


Fig. 6: Comparison of distribution pies (observed actions)[5]

There are obvious similarities between the journalists from FAZ and Netzeitung. The distribution pies for the editors of the Netzeitung actually look as if they were Siamese twins. So both of them are doing almost the same things and spend approximately the same amount of time on similar actions, although they clearly are two different individuals.

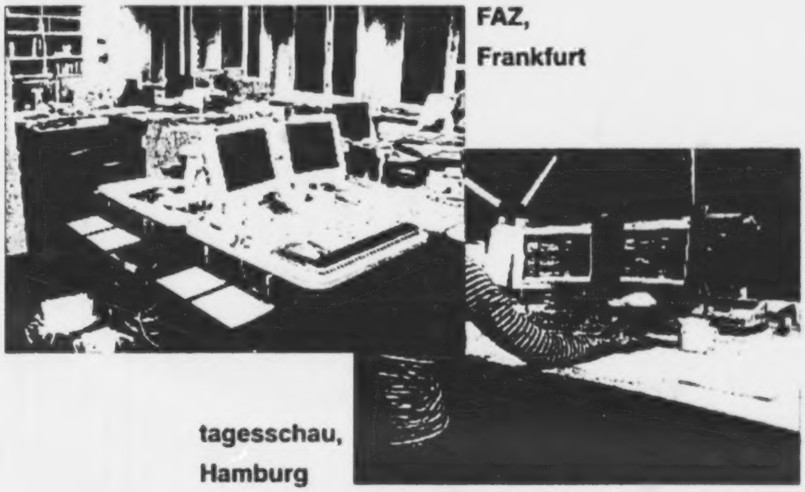


Fig. 7: Workplaces of online journalists

Another example: The above photographs (cf. fig. 7) depict the workplace of two online journalists. They look very similar: A lot of printouts, and two flat screens. The journalists used the two screens for just the same reasons (content management system on one screen, agency news on the other). It is worth noting that in both cases, the flat screens were bought by the media companies because the journalists asked their management to do so. So that was not a structure that shaped the journalists action in the first place, but they would ask for this setup due to their working necessities. The most surprising fact however is that the pictures were taken at two rather different media, namely the FAZ in Frankfurt, and the Tagesschau in Hamburg. The main company of Tagesschau online is a public TV station like the BBC, while the main company of the FAZ is a conservative nationwide newspaper.

---

[5] All numbers have been omitted for the sake of clarity.

So we conclude that some individuals in the observed newsrooms have obviously developed comparable working patterns, and that they are using comparable resources in similar working places.

4.2.2 Associations and sequences

According to the above-mentioned theory, some actions refer to certain resources, resources to places or time frames, time frames to actions and so on. These relations can be described by associations and sequences contained in the data matrix. With the help of the standard data mining program Clementine, we carried out a network analysis of these associations. In principle, such an analysis counts the connections between the individual values of the coded variables and compares the actual number of observed connections between two values with the overall number of connections of the first value. Therefore it gives us an overall impression of the strongest connections (for example, it will give you an impression about the strength of the ties between certain actions and resources). The network viewgraph shown below contains all the action types (on the left) and resources (on the right) that were observed during our study (cf. fig. 8).

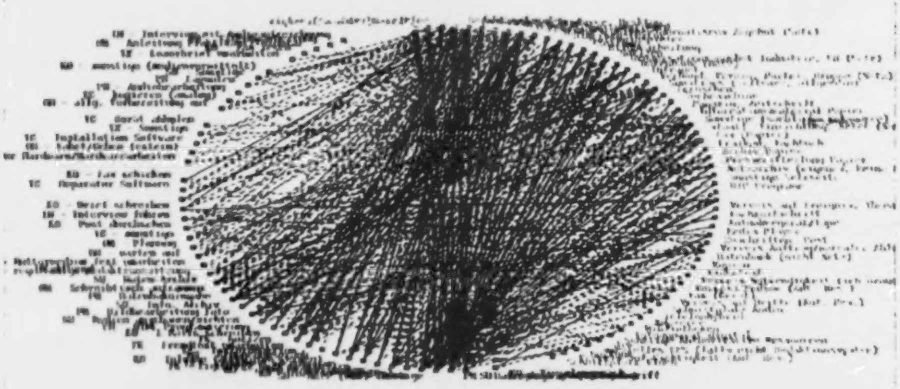


Fig. 8: Association analysis (action type x resources)

Obviously, these connections are not evenly distributed. There are some strong ties, and a lot of weak ties – and quite a lot of nodes are not connected at all. A change of threshold within Clementine's network analysis algorithm will highlight the most frequent connections and delete all of the weaker ties (cf. fig 9).

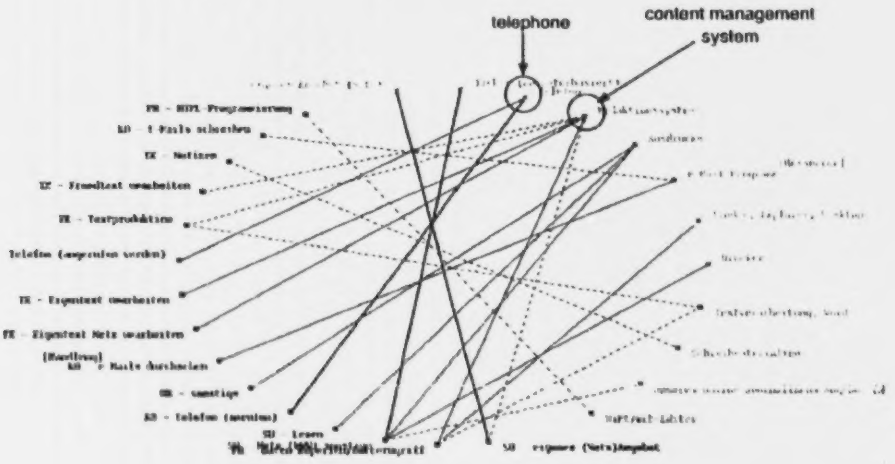

telephone
content management system
PR - HTML-Programmierung
KO - E-Mails schreiben
TK - Notizen
TE - Fremdtext bearbeiten
TR - Textproduktion
Telefon (angerufen werden)
TK - Eigentext bearbeiten
FK - Eigentext Netz bearbeiten
[Handlung]
KO - E-Mails durchsehen
SO - sonstige
KO - Telefon (anrufen)
SU - Lesen

Fig. 9: Association analysis, strongest connections (action type x resources)

As the number of connections is shrinking, one is finally left with just the strongest ties – which are of course very obvious links, after all. For example, the communication acts are very strongly related to the resource 'telephone'. That hardly comes as a surprise. Nevertheless, it is also a clear indication that the resources were well defined, that that they just serve one major purpose. Content management systems, on the other hand, are of a rather different nature. They are used as central nodes for many types of actions. This may indicate the importance of such tools for the production of news in online journalism. It is also obvious from the network diagram that action types always refer to the same arrangement of resources (one action leads to one resource which leads to another action etc.). These relations create robust sense-making patterns, because resources and actions are really glued together by such links.

Another type of analysis focuses on the temporal sequences of individual acts over time. While this can be carried out with the help of sequence analysis algorithms as well (like Clementine's CAPRI algorithm), we chose to carry out a graphical analysis first. As Keim (2002; see also Klösgen & Zytkow, 2002, 226 ff.) notes, graphical analysis by a human being can be superior to computer algorithms, simply because humans easily detect certain patterns on the basis of their huge knowledge of similar observed phenomena. The granulation of the observation is a difficult problem for computer programs ("what is the size of the elements that should be observed, how long should the sequences be etc.?"). Similar problems appear when it comes to the interpretation of raw data ("what kind of sequence is trivial, what kind of sequence is important?").

In order to analyze those sequences, we applied a "slicing" algorithm to the material, cutting the observed actions into 5-second pieces (the starting point and the end point of each action have been coded). The resulting data consisted of temporal cases, where each case represents an equal amount of time. Based on this transformed data set, it was possible to produce a graphical

BEST COPY AVAILABLE

display of actions over time that may form the basis for further analysis (cf. fig 10 for an example).[6]

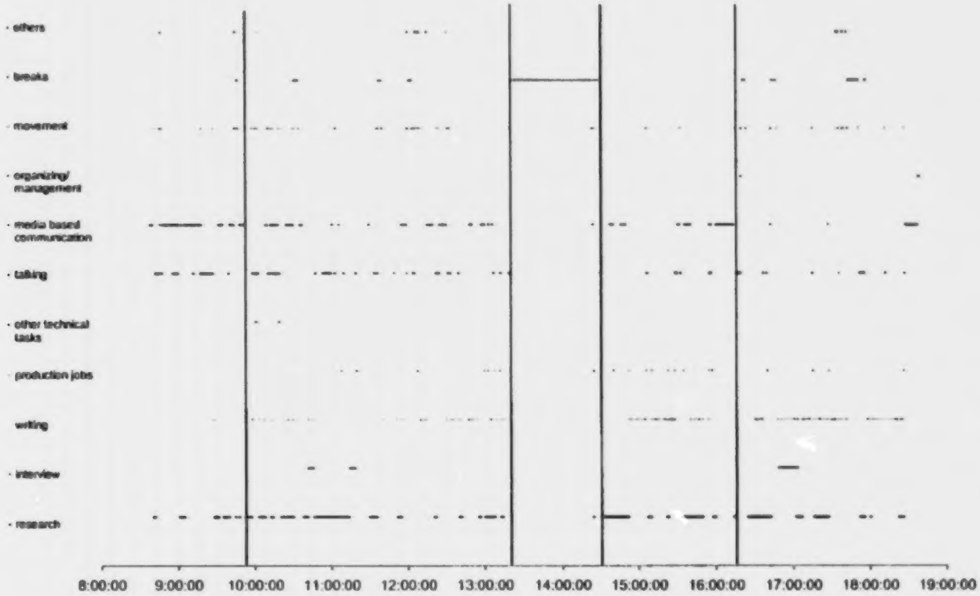


Fig. 10: Time based graphical analysis, "piano roll" graph (for action types)

The picture represents one working day. Each row shows one category of actions. The vertical lines depict the starting/end points of different phases of a working day: In the beginning (8.30 - 9.50), the online journalist is reading a lot of e-Mails (media based communication) and does not write very much. Which is something like an orientational phase that marks the beginning of almost every working day. The following bigger "work phase" shows quite a lot of research in the beginning (which also serves a first orientation to collect interesting news), with more writing during the second half of this period. Then there is a break (13.30 - 14.30). After the break, there is a second working phase with almost no movements, but phone calls, e-mail exchanges and long sequences of research. The last period of the day is characterized by writing and researching articles (those actions are usually bound to one news article/topic). There is almost no phoning/media based communication going on (only a few contact persons were available after 16:00, although there was a big interview taking place afterwards, in this special case), and this period is dominated by long sequences of writing .

This is quite a common pattern for online journalism. Obviously there is a constant stream of writing and research happening during the working day. There are no real production deadlines, but a constant need for researching and reworking news. Nevertheless, some of the communication processes seem to fade out by the end of a working day, which is dominated by writing.

---

[6] One of the biggest problems was the possibility of multiple actions taking place at the same point of time. But the „slicing" of the data would allow for the transformation of an action-based matrix (1 case = 1 action) into a time-based matrix (1 case = 1 time step, with new variables describing all the actions at this point of time).

And despite the fact that there are no real production deadlines, we still observe orientational phases at the beginning of each working day and production peaks during the day. External contacts also seem to shape working patterns to a certain extent.

Without going into the details, it may be noted that this finding is in clear contrast to various speculations which claim that online journalism may not be bound to the restriction of time.

A detailed analysis of associations and sequences among our data (remember that we can go down to details of five seconds) also shows that there are interesting work patterns which seem to develop into rules of action. For example, writing and research routines seem to follow similar patterns in the vast majority of all cases, with a consistent use of content management systems and satellite/internet based news agency information as basic working resources.

This may illustrate some of the possibilities of the theoretical and empirical approach. In a last section, we will summarize the pros and cons of the network perspective, and draw some final conclusions.

5. Conclusion: Theoretical and empirical potential

When using a new approach, one has to ask: What might be the advantages, what might be the disadvantages of such an approach, compared to traditional ways of looking at society and action? In that sense, there are certainly some disadvantages to the network perspective:

- The theoretical approach is only loosely tied to traditional perspectives about journalism.
- It offers no simple, singular description of phenomena in the media.
- It depicts complicated, fluctuating relations that may lead to ambiguities and sometimes even to contradictions.

Nevertheless, there are also advantages to such a procedure:

- It takes into account the complexity of the social construction of reality.
- As an analytical approach to the production of action networks and meaning in every day action (similar relations of elements), it gives us more than just a handful of metaphors for describing social phenomena.
- It is an inherently dynamic view, which is helpful if you want to look at changing aspects of social life.
- It is open for empirical research.

Empirically, the network-based observation of human action shows its potential when it comes to a detailed description of working behavior. It can certainly be carried out in addition to surveys, as a supplement and to correct certain aspects. One has to note however that observational studies cannot be carried out on a representative basis, because they are just too costly and would interfere with the editorial processes if carried out on a large scale. But they can be employed to create empirical conjectures, uncovering unknown relations with the help of data mining tools.

These tools have many applications beyond the few examples shown in this paper. Algorithms based on (neural) networks (Klösgen & Zytkow, 2002) could lead to a deeper understanding of human behavior. We believe that for journalism research, this could open up a new way of theoretical thinking as well as new ways of empirical research and analysis.

References

(Translations in parentheses)

Barabási, A.-L. (2002). Linked. The New Science of Networks. Cambridge: Perseus.

Biggs, N. L., Lloyd, E. K., & Wilson, R. J. (1976). Graph Theory: 1736-1936. Oxford: Clanderon Press.

Castells, M. (2000). The Rise of the Network Society. Oxford: Blackwell Publishers.

Giddens, A. (1997). Die Konstitution der Gesellschaft. Grundzüge einer Theorie der Strukturierung (The Constitution of Society. Outline of the Theory of Structuration). Frankfurt a.M., New York: Campus.

Keim, D. A. (2002). Data Mining mit bloßem Auge (Data Mining with the Eyes Only). Spektrum der Wissenschaft, November, 88-91.

Klösgen, W., & Zytkow, J. (Eds.). (2002). Handbook of Data Mining and Knowledge Discovery. Oxford: Oxford University Press.

Neuberger, C. (2000). Renaissance oder Niedergang des Journalismus? Ein Forschungsüberblick zum Online-Journalismus (Renaissance or Decline of Journalism? An Overview of Research on Online Journalism). In K.-D. Altmeppen, H.-J. Bucher & M. Löffelholz (Eds.), Online-Journalismus. Perspektiven für Wissenschaft und Praxis (Online jouranlism. Perspective for science and practical applications) (pp. 15-48). Wiesbaden: Westdeutscher Verlag.

Perrow, C. (1992). Small-Firm Networks. In G. Eccles Robert & N. Nohria (Eds.), Networks and Organizations: Structure, Form, and Action (pp. 445-470). Boston: Harvard Business School Press.

Quandt, T. (2003). Journalisten im Netz. Eine Untersuchung journalistischen Handelns in Online-Redaktionen (Journalism on the Internet. Journalistic Action in Online Newsrooms. A Study). Technische Universität Ilmenau, Ilmenau.

Schütz, A. (1981). Der sinnhafte Aufbau der sozialen Welt (The Phenomenology of the Social World) (2. ed., first original ed. 1932). Frankfurt a.M.: Suhrkamp.

Schütz, A. (2002). Theorie der Lebenswelt (The Structures of the Life-World) (New ed.). Konstanz: UVK.

Scott, J. (2000). Social Network Analysis. A Handbook (2. ed.). London: Sage.

Scott, J. (Ed.). (2002). Social Networks: Critical Concepts in Sociology. London, New York: Routledge.

Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

Weber, M. (1972). Wirtschaft und Gesellschaft (Economy and society) (5., revised ed., original ed. 1921/1922). Tübingen: Mohr.

Williamson, O. E. (1985). The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. New York , London: Free Press : Collier Macmillan Publishers.

Windeler, A. (2001). Unternehmungsnetzwerke. Konstitution und Strukturation (Organizational Networks. Constitution and Structuration). Wiesbaden: Westdeutscher Verlag.

COMMUNITY & CIVIC VALUES, COMMUNICATION, AND SOCIAL CAPITAL

"Bowling Alone" as a Product of Values and Communication

By
Leo W. Jeffres, Jae-won Lee, Kimberly Neuendorf, and David Atkin
Department of Communication
Cleveland State University
Cleveland, Ohio 44115

Correspondence:
address to Leo W. Jeffres
phone (216-687-5088)
email: l.jeffres@csuohio.edu

Paper presented to the Communication Theory & Methodology Division for presentation at the annual conference of the Association for Education in Journalism and Mass Communication, Kansas City, Mo., July 30-Aug. 2, 2003.

Abstract

Community & Civic Values, Communication, and Social Capital: "Bowling Alone" as a Product of Values and Communication

When Putnam (1995) focused attention on a decline in organizational involvement, he renewed interest in community activities and their consequences for civic life. Since civic involvement occurs at the most "local" level, the community and neighborhood have emerged as contexts for examining "social capital" and processes involved in its decline or direction. This paper examines relationships between civic and community values, communication variables and community variables that include social capital, community attachment and identity, using data from a survey conducted in a Midwest metro area in the summer of 2001. Values associated with civic culture, social networks and social capital are strongly related to community attachment, organizational ties and quality of life assessments. The only media use variable consistently related to such values is newspaper readership, which is positively correlated with values representing civic culture, social networks and social capital. Reading the newspaper also is positively correlated with measures of community attachment, community activities, belonging to organizations and both metro and neighborhood quality of life assessments; however, going out to see films, using other print media, and frequent use of the Internet also are important. Interpersonal communication variables are related to both community variables and civic and community values. Regression analyses predicting a summary measure of social capital show that communication variables explain additional variance beyond that accounted for by social categories, civic values and community values.

Community & Civic Values, Communication, and Social Capital: "Bowling Alone" as a Product of Values and Communication

When Putnam (1995, 1996) focused attention on a decline in organizational involvement, he renewed interest in community activities and their consequences for civic life, particularly political participation. As de Tocqueville (1848/1969) argued, American democracy is anchored in its civic associations, which Putnam (1993) called "fabrics of trust."[1] Such associations are thought to generate ideas and develop understanding, producing civic engagement. For communication scholars, the emphasis has been on how communication variables are woven into models predicting political participation. Here we shift attention to how communication and values fit into a network of predictors of community involvement.[2]

Putnam pointed to evidence suggesting that participation in such organizations as formal clubs and organizations, unions and committees has steadily declined over the past three decades (Putnam & Yonish, 1997),[3] and levels of interpersonal trust have accompanied these declines. Putnam (1995b) sees these factors as mutually causal and identified television as one of the chief culprits responsible for this decline in "social capital." Shah et al. (1999) added the notion of life satisfaction, or contentment, as an intrapersonal dimension and then tested the relationships posited by Putnam. They found strong reciprocal relationships among civic engagement,[4] contentment or satisfaction with one's personal life, and interpersonal trust.[5] However, although total television viewing was negatively related to involvement in collective activities, public affairs media use was positively related to civic engagement. Thus, the role of the media is not solely negative, and consequences of exposure depend on uses and gratifications as well as the medium (Shah, McLeod & Yoon, 2001).[6] Shah, McLeod and Yoon (2001) show the strong ties between informational uses of the media and "social capital," measured as volunteer work, attending organizational meetings and working on projects. Among younger adults, use of the Internet for exchanging information strongly influenced trust in people and civic participation more than did uses of traditional print and broadcast news media.

Others have tested media use patterns and the other key variable in Putnam's scenario, trust in public institutions.[7] Moy, Pfau and Kahlor (1999) found that viewing news on television predicted positive perceptions of the news media and public schools, while newspaper reading was positively correlated with favorable attitudes towards the criminal court system and schools. Pfau, Moy, Radler and Bridgeman (1998) found a negative pattern of relationships between public perceptions of democratic institutions—the presidency, Congress, news media and public schools—and viewing network television news, other television news, television entertainment talk shows and political talk radio shows. Moy, Scheufele and Holbert (1999) tested Putnam's (1995a, 1995b) charge that television is the force behind the decline in social capital in America because it inhibits participation outside the home. In their survey, time spent with television did not affect civic engagement through perceptions of time pressure, but there was a direct negative impact of television viewing time on civic engagement. However, time spent reading newspapers enhanced engagement, and the more time spent with newspaper, the less time pressures respondents perceived. Individual differences also are a factor, particularly for the new technologies.

The Community Context:

Since civic involvement occurs at the most "local" level, the community and neighborhood have emerged as contexts for examining "social capital" and processes involved in its decline or direction. Forrest and Kearns (2001) note that the neighborhood again has gained prominence as a setting for processes that shape social identity. They examine debates around the concept of social capital and break it down into domains for policy action at the neighborhood level: empowerment, participation, associational activity and common purpose, supporting networks and reciprocity, collective norms and values, trust, safety and belonging.

In urban centers as well as small rural communities facing economic difficulties, citizen involvement is important for the success of development. Orr and West (2002) look at citizens' views on urban revitalization and conclude that it is a multidimensional phenomenon. Docherty, Goodlad and Paddison (2001) use qualitative evidence from four neighborhoods in two cities, concluding that civic participation in urban governance is fostered by political structures and public policy as well as a civic culture supportive of citizen involvement. Since the neighborhood movement began in the early 1960s, it has evolved through different stages. Keyes (1987-1988) notes that the major goals in the early 1960s were concern for turf, control of area decisions and resident participation, with communication and access to resources gaining prominence in the 1980s.

The relationship between communication and community ties is not of recent vintage, but it's important for the discussion about social capital. Community ties include involvement in organizations but the concept also is used in a broader sense. Community ties also are used as indicators of community integration, and the extent to which residents become involved with and attached to their communities (see Demers & Viswanath 1999; Jeffres, 2002; Stamm, 1985). Stamm, Emig and Hesse (1997) note that some media contribute to community integration more than do others. Finnegan and Viswanath (1988) found that regular reading of the local community weekly was correlated with neighborhood involvement and with use of local community facilities. Jeffres and Dobos (1984) found that attention to neighborhood newspapers led to awareness of local groups, while interpersonal communication was closely linked to participation in neighborhood projects. Also, education was correlated with awareness, and income was correlated with participation in self-help projects. Emig (1995) looks at the relationship between community ties such as voting in local elections and keeping up with local news, and use of certain types of media. The study found that media use varied according to the types of community ties exhibited by individuals. Use of newspapers to follow what's going on locally in the community was linked to being a registered voter and voting in the last election.

Interpersonal communication is as important as mass communication to discussions of community ties. The two modes of communication also reinforce each other, as media use provides topics for subsequent discussions, and arguments that arise lead to subsequent media use. Scheufele (2002) points out that interpersonal communication about politics and public affairs has long been treated as the "soul of democracy" but the relationship between hard news media use and political participation cloaks different effects for people who talk about politics frequently versus those who do not. McLeod, Daily, Guo, Eveland, Bayer, Yang and Wang (1996) looked at community integration, media use and democratic processes, using 15 indicators drawn from integration studies to test the hypothesis that community integration is a multi-dimensional concept. Psychological attachment to one's community was correlated with the strength of one's interpersonal network and with an emphasis on localism versus cosmopolitanism.[8] Geiger, Bruning and Harwood (2001) found that people talk about news and prime-time television programs more than they do about other types of programs. Older adults were more likely to talk about such highbrow programs as those on public broadcasting and news programs, while younger adults talked more about such programs as soap operas, animation, and science fiction. Discussion about news programs focused on issues. Other processing variables clearly are important, as people integrate information and perceptions from different sources.[9]

At the community level, citizen involvement and activity are often linked to perceived quality of life issues as well as community attachment. Jeffres and Dobos (1995) found that satisfaction with life was positively correlated with interpersonal measures of frequency of talk with others about problems in the area and frequency of communication with neighbors. Doolittle and MacDonald (1978) look at the relationship between communication and sense of community in a stable Milwaukee neighborhood, finding that residents communicated easily and frequently among themselves, were aware of numerous interaction possibilities, and demonstrated skills in the use of internal and external communication systems for a variety of purposes, including influencing agencies, officials and institutions outside of the neighborhood. Ball-Rokeach, Kim and Matei (2001) offer a model in which neighborhood storytelling is the communication process through which people become members of a neighborhood. Thus, those who are actively involved in such interpersonal communication had a stronger sense of belonging to the neighborhood.

In neighborhoods and communities, the relationship between organizations and communication is important for development projects as well as more enduring factors such as sense of community. O'Hara (2001) cites ineffective communication between local residents and decision makers as one of two barriers that keep urban neighborhoods isolated from the larger environment. Peyrot and Fenzel (1994) note that neighborhood organizations serve as a mechanism for diffusing information and mobilizing residents through information gathering and establishing communication channels. Hyland and Ciaramitaro (1984) describe a community intervention model in energy conservation in Memphis, Tenn., where new modes of communication and empowerment of low-income residents contributed to increased participation in an energy conservation program. Becker and Wehner (1998) also suggest that interactive electronic media will be useful to non-governmental organizations, community pressure groups and local activities.

Civic Values:

Accompanying Putnam's (1995, 1996) concern over a decline in "social capital" is the idea that societies face a new crisis of social cohesion. A measure of consensus is necessary for the social processes needed to make democracy work at all levels. Forrest and Kearns (2001) suggest that the neighborhood as a source of social identity is being eroded with the emergence of a more fluid, individualized way of life. They also identify five domains of social cohesion: common values and a civic culture (e.g., common aims and objectives, support for political institutions and participation in politics); social order and social control (e.g., absence of general conflict and threats to existing order, tolerance, respect for difference); social solidarity and reductions in wealth disparities (e.g., equal assess to services and benefits, acknowledgement of social obligations and willingness to assist others); social networks and social capital (e.g., high social interaction within community, civic engagement and associational activity); and place attachment and identity.[10]

McCombs (1997) notes that the news media generate community consensus by framing issues, giving some issues more prominence than others, and focusing community groups on particular issues, the agenda-setting function of the media. However, media audiences themselves also hold a set of values about their governments and community, and their involvement in each. The research reviewed above shows support for relationships between media use, interpersonal communication (networks, relationships and topics), community ties and attachments, involvement in communities and organizations, political involvement and confidence/trust in people and democratic institutions. While the values people hold are cited for their importance, they have not been given a position in models tested with empirical data.

In their effort to identify the domains of social capital, Forrest and Kearns (2001) cite the key dimensions of social cohesion, which include three that refer to common values, one that combines communication and organizational involvement, and one that specifies community identity. We will investigate the relationships among these concepts using data drawn from a larger study on people's values.

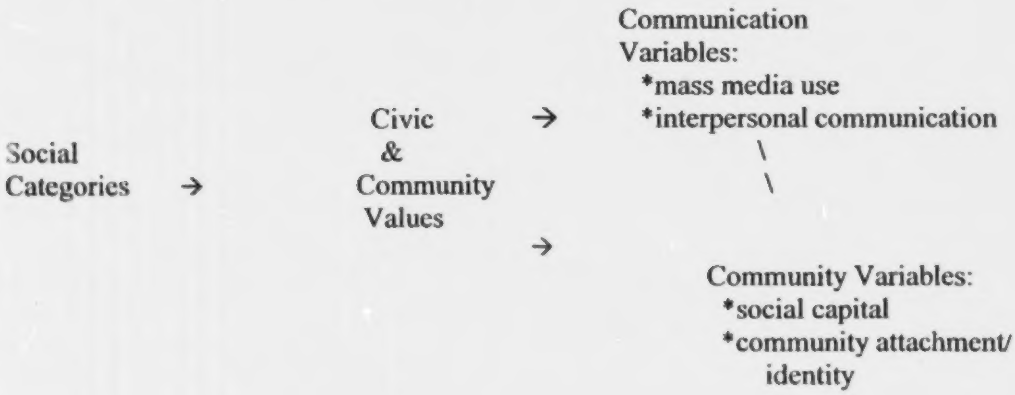
A Model for Social Capital at the Community Level— Values, Communication, and Identity

The five dimensions of social cohesion identified by Forrest and Kearns (2001) include several that focus our attention on values: 1) a civic culture (common aims and objectives, support for political institutions; 2) social order and social control (tolerance, respect for differences and an absence of general conflict and threats to existing order,); 3) social solidarity and reduced wealth disparities (equal assess to services and benefits, acknowledgement of social obligations and willingness to assist others); and 4) social capital values (community interaction and civic engagement). Number 1 is the first amendment of social cohesion—support for the system itself, whether we agree with current policies or office holders. The second underlines the importance of tolerance, or respect for differences of opinion; in a democratic system, influence is brought to bear peacefully, not through force of arms. Competition is through communication and not through coercion. The third stresses the importance for notions of equal access, acceptance of one's role as a citizen, and a willingness to help others. Although these may fluctuate across time, they should be relatively enduring for people as individuals, with changes likely to occur during times of external threats, e.g., wartime, or internal system crises, e.g., Watergate. The fourth is social capital, here viewed as the extent to which individuals value their community and civic ties.

Several of the key dimensions also include behavioral components in addition to the values. Thus, the fourth factor includes social networks—through interpersonal channels as well as mass communication—and civic engagement, the latter including involvement in associations. The fifth dimension again is a combination of attachment and identity, i.e., one feels attached to one's community and derives a certain amount of identity from it.

We propose to examine relationships among the different sets of variables represented in Figure 1: civic and community values; community variables representing social capital and community attachment/identity; and communication. This three-legged stool is preceded by social categories that represent the individual differences affecting each group or relationships among them.

Figure 1: A Model Relating Civic & Community Values to Social Capital and Communication Variables


Social Categories
→
Civic & Community Values
→
→
Communication Variables:
*mass media use
*interpersonal communication
\
\
Community Variables:
*social capital
*community attachment/identity

After an examination of the inter-relationships, we will investigate the strength of social categories, civic and community values and communication variables as significant predictors of social capital and the other community variables.

Methods

A survey was conducted from June 20-July 8, 2001 in a major Midwest metropolitan area, using a random sample of residents and interviews conducted with a CATI (computer-aided telephone interviewing) system. Telephone numbers were selected through random-digit dialing procedures.[11] The survey was presented as a general poll with an emphasis on values and what people think is important in life. The response rate was about 45 percent. Variables used in this paper were operationalized as follows:

Importance of Values—The key values for this study were imbedded in a list of 30 values rated in terms of their importance. Procedures followed those used by Tan et al. (1997) we asked respondents to use a 0 to 10 scale to "rate how important each of the following values are to you personally," where "0 means it's totally unimportant, 10 means it's extremely important and 5 is neutral."

> **Civic Culture Values**—Respondents were asked to rate the importance of "participating in the political system" and "being a good American" using a 0-10 scale. The two values are correlated at .13 (p<.05).
>
> **Social Order & Control Values**—Respondents were asked to rate the importance of "tolerance of other people" and "being obedient to authority" using a 0-10 scale. The two values are correlated at .21 (p<.001).
>
> **Social Solidarity Values**—Respondents were asked to rate the importance of "equality for all" and "being able to help others" using a 0-10 scale. The two variables are correlated at .45 (p<.001).
>
> **Social Networks and Social Capital Values**—Respondents were asked to rate the importance of "being involved in the community" and "having good neighbors" using a 0-10 scale. The two values are correlated at .37 (p<.001).

Community Variables—Several measures were constructed for community attachment, community activities, organizational ties, and community assessment. The operationalizations were based on items used in other studies (see Author).

> **Community Attachment/Identity**—Respondents were asked how much they agreed with the following statement, "I feel a strong attachment to my community." Respondents were directed to use a 0-10 scale where 0 meant they strongly disagree, 5 was neutral and 10 meant they strongly agreed. This item is similar to those in Fessler's (1952) community solidarity index (see Miller, 1991).
>
> **Community Activities**–Respondents were asked to use a 0-10 scale to tell how often they did each of several things, including going to sporting events, going to cultural events such as plays or the orchestra, going to local museums, attending concerts of current musical groups or artists. Responses to each item were standardized and the scores summed up for a scale (alpha = .68).

Organizational Ties—Respondents were asked, "Do you belong to any neighborhood or community organizations, including block clubs, social groups, religious groups, business groups or ethnic groups?" If they said yes, they were asked, "What are they?" The number cited was coded. This item has been used in various studies cited above.

Community Assessment—The first two items in the interview schedule asked respondents to assess the quality of life in the metro area and their neighborhood, each using a 0-10 scale where 0 means the "worst place to live" and 10 means the "best place to live." Responses were examined individually and standardized and added for a scale of community assessment. The two assessments are correlated at .34 ($p<.01$). Both measures are standard items from the quality of life literature (e.g., Andrews, 1986).

Media Use--Media use was measured using the standard set of items as well as several measures for the new technologies:

TV Viewing—Respondents were asked for the number of hours of television they watched "yesterday." The scale ranged from 0 for none to 11 for more than ten hours.

TV News Viewing—Respondents were asked how often usually watch the news on television, several times a day, about once a day, 5 or 6 days a week, 3 or 4 days a week, 1 or 2 days a week, or less often than that.

Radio Listening—Respondents were asked how many hours they listened to the radio yesterday. Coding was done using the same scale used for television.

Newspaper Reading—Respondents were asked how many days last week they read a newspaper, and responses were coded from 0 to 7.

Magazine Reading—Respondents were asked how many different magazines they read regularly. Responses were coded into 8 categories, 0, 1, 2, 3, 4, 5, 5 to 10, 11 to 20, and 21 or more.

Book Reading—Respondents were asked how many books they read in the past six months. Responses were coded into the same 8 categories used for magazines.

Video Viewing—Respondents were asked how many borrowed or rented videos they watched in the past month. Responses were coded into the same 8 categories used for magazines.

Film Viewing—Respondents were asked how many times they went out to see a movie in a theater in the past month. Responses were coded into the same 8 categories used for magazines

Media Use Index—Responses to the use of traditional media were standardized and the scores summed up for an index.

Computer Access—Respondents were asked if they had a personal or laptop computer in their household and responses were coded yes or no.

Internet Access—Respondents were asked if they had access to the Internet at home, at work, or both. Access was coded two ways, as a dummy variable where access anywhere =1 and no access = 0; and as a continuum where access at both home and work =2, access at either alone =1, and no access = 0.

Internet Use—Respondents were asked if they had ever gone on the Internet. Those who said yes were asked how often they go on the Internet at work, using a 0-7 scale ranging from several times a day to almost never. They also were asked how often they go on the Internet at work using the same scale. Several variables were constructed: 1) a simple usage measure where 1=has gone on the Internet before, 0=has never gone on the Internet; 2) Internet access (access = 1); 3) Frequency use Internet at work (those without access = -1); 4) Frequency use Internet at home (those without access =-1); 5) Overall Internet Use combining the scores for usage at home and work.

Media Website Use—Respondents were asked how often they visited media websites such as one of the TV networks, a newspaper or radio site, using a 7-point scale ranging from almost never (1) to several times a day (7). Those not using the Internet were assigned a value of 0.

Chat Room Use—Respondents were asked if they had ever visited a chat room on the Internet to talk with people about something. Those who said yes were asked how often, using the following categories: every day (6), a couple times a week, about once a week, a couple times a month, less often than that (2). Those who had never visited a chatroom before were assigned a 1 and those who had never gone on the Internet were assigned a 0.

Social Categories--The standard social categories were measured, including: marital status, the number of people in one's household, age, level of formal education completed, ethnic or racial background, household income and gender. Dummy variables were constructed for being married, being white, being black, and being other race or ethnicity.

Interpersonal Communication Relationships— Two measures of interpersonal communication relationships were used. Respondents were asked how much they agreed with each of the following statements, using a 0-10 scale where 0 meant they strongly disagree, 5 was neutral and 10 meant they strongly agreed: "I enjoy meeting strangers" and "I prefer talking with people who have the same background as me."

Interpersonal Communication about Public Affairs/Media Topics-- Respondents were asked how often they did each of several things using a 0-10 scale where 0 meant they never did this and 10 meant they did this all the time: "talking about current events with friends or coworkers" and "talking about things I've seen in the media with friends or coworkers."

Results

First relationships between civic and community values, and community are analyzed. As Table 1 shows, one of the two values representing civic culture—participating in the political system—is strongly related to all of the community variables, and the relationships persist when social categories are controlled (ascriptive—gender, ethnicity; status—education, income; life cycle—married, age, household size). The other value, being a good American, is correlated with community attachment and assessments of the neighborhood quality of life but not with the other community variables.

Values representing social order and control are positively related only to quality of life assessments. Tolerance of other people has no relationship with community attachment or organizational ties, key variables in the social capital literature. This value also is unrelated to participating in community leisure activities. The other value, being obedient to authority, shows a mixed pattern of relationships. The more important obedience is, the stronger community attachment, the higher the neighborhood quality of life assessment and the less involved in community leisure activities, but these relationships drop below statistical significance with controls. Organizational ties are unrelated to either social order and control value.

The two social solidarity values are positively related to three community variables but not organizational ties. Thus, the more important equality is as a value, the stronger one's community attachment and the higher one's metro and neighborhood quality of life assessments; the relationship with community attachments falls below statistical significance with controls. Being able to help others is correlated with community attachment but with none of the other variables, including participation through organizations.

The strongest pattern of positive relationships in the table is found with values representing the social network and social capital values themselves. Thus, the more important being involved in the community is as a value, the stronger the community attachment, the more community leisure activities one is involved in, the more organizations one joins, and the stronger the metro quality of life assessment. Only the neighborhood quality of life assessment is unrelated. Having good neighbors is correlated with community attachment, metro and neighborhood quality of life assessments and belonging to organizations, but not to attending community leisure activities, which occur at a metro-wide level rather than the neighborhood.

Next relationships between values and traditional media use variables are examined (see Table 2). In general, there are few consistent relationships across variables, with one exception—reading the newspaper. Consistent with results from other studies examining political behaviors, we find that reading the newspaper more frequently is correlated with both civic culture values, with one of the two social order and control values (tolerance but not being obedient) and with the first social network and social capital value, being involved in the community; these relationships persist with controls. The correlations between being a good American and the two measures of television viewing (hours and news) approach significance but drop out with controls. There also are a couple negative correlations between going out to see films and values but these too decline with controls.

Thus, with values affecting community variables but only one major media use variable, we move on to examine relationships between the latter two sets of variables (see Table 3). Here we see a strong pattern of positive relationships between reading the newspaper and community variables—community attachment, attending community leisure activities, belonging to community organizations and quality of life assessments, although the last of these declines with controls. In general, use of the other print media also is correlated with attending community leisure activities and belonging to organizations, relationships that persist with controls. Going out frequently to see films in a theater also is positively correlated with attending other community leisure activities but is negatively correlated with community attachment.

The new communication technologies offer another opportunity for people to connect with their communities. Table 4 shows that few of the measures of Internet use are correlated with either community attachment or belonging to organizations, but most are correlated with attending community leisure events. Thus, the Internet joins some traditional media in promoting such community involvement.

Table 5 shows relationships between the measures tapping Internet use and community and civic values. We see a mixed pattern where the isolated relationships that appear are generally negative. For example, those for whom being a good American is important are less likely to go on the Internet or to use it frequently. The overall Internet use measure—which reflects access, whether one has gone on the Internet and overall current use—is negatively correlated with the importance of being a good American, being obedient to authority, and being able to help others. Although relationships decline with controls, the pattern suggests that those most with the strongest community links are not on the Internet or that the Internet does not promote these values. A static data set cannot differentiate the direction of influence.

Interpersonal communication variables are related to both community variables and civic and community values in Table 6. Our measures include two tapping "relationships," e.g., the extent one enjoys talking to strangers, and two that link to mass communication, i.e., talking frequently about things in the media or current events. As the table shows, those who enjoy meeting strangers say that three sets of values are more important—social order and control, social solidarity, and social network and social capital. Thus, those for whom tolerance, obedience, being able to help others, being involved in the community and having good neighbors are important, also enjoy meeting strangers. We also see that the same values generally are associated with talking more frequently about current events or other things from the media. The only set of values unrelated to the interpersonal communication measures are civic culture values.

Community variables of attachment, organizational ties and quality of life assessments also are correlated with the interpersonal measures. Thus, those who enjoy meeting strangers are more strongly attached to their community, belong to more organizations—they're not bowling alone, and have higher metro and neighborhood quality of life assessments. Those who prefer talking to people like themselves also are more strongly attached to their community and assess their neighborhood quality of life more highly but they are not involved in more organizations and do not attend more community leisure activities. And there is an almost complete pattern of positive correlations between talking about current events or things in the media and the community variables. Those who engage in such conversations more often are more attached to the community, attend more community leisure events, belong to more organizations, and assess the quality of life more highly. Most of these relationships persist with controls.

Predicting Social Capital

Although questions of causality cannot be addressed with a static data set, we can examine the extent to which social categories, sets of civic and community values, and communication variables predict social capital and related community variables of attachment, activities and assessments. Table 7 shows the results of regressions predicting each of these factors. The community measures—community attachment, organizational ties, quality of life assessments and community activities were factor analyzed and one factor emerged, with loadings ranging from .59 to .70. The factor accounted for 43 percent of the shared variance. Respondents factor scores were used as a **summary measure for social capital**.

As Table 7 shows, both values and communication variables are significant predictors for the factor capturing respondents' loadings for the social capital factor dimension. In addition, values are key predictors for organizational ties, community attachments and quality of life assessments. The most significant values are participating in the political system and being involved in the community; the former is a significant predictor for three dependent measures—the Social Capital factor score, organizational ties, and quality of life assessments. The latter—being involved in the community—is a significant predictor for the Social Capital factor score, organizational ties, community attachment and community activities. Having good neighbors was a significant predictor for the Social Capital factor score and quality-of-life assessments. Equality for all was a significant predictor for quality of life assessments. None of the other values were significant predictors.

Communication variables are key predictors for the overall social capital score, as well as for measures of community attachment, community activities and quality-of-life assessments. In only a few instances were media use variables significant predictors; newspaper use, going out to see films and visiting chat rooms predicted attending community leisure activities, while listening to the radio and Internet use predicted community attachment. Interpersonal relations measures were significant predictors in a couple instances; enjoying meeting strangers was a significant predictor for the Social Capital factor score, for organizational ties, and for quality-of-life assessments, while a preference for talking with people of the same background was a significant predictor of the Social Capital factor score, community attachment and the quality-of-life assessment. The frequency with which one talks about things in the media was a predictor of the Social Capital factor score, community attachment, and community activities. The frequency with which one talks about current events was a significant predictor for quality-of-life assessments.

Social categories are significant predictors for community activities and quality of life assessments. Education was a significant predictor for organizational ties and community activities, and age was a predictor for community attachment. Married status was a negative predictor for community activities. Age and ethnicity were significant predictors for quality-of-life assessments and the standardized beta for household income approached statistical significance.

Discussion

Putnam (1995, 1996) focused on organizational ties, using the metaphor of "bowling alone." That activity here is represented in organizational ties—actually belonging to community groups--and in community activities—going to sporting events, local museums, concerts and similar events. Neither of those community variables is related to social solidarity values, or social order and control values. Interestingly, those who value having good neighbors more highly are not much more involved in their communities. However, other aspects of social capital, including community attachment and quality of life assessments show a stronger pattern of correlation with different civic values.

In a variety of research traditions, media have been seen as having negative effects on social capital and other sources of support for the political system. In the 1970s, media treatment of government was seen as creating a "videomalaise" represented by disillusionment with government and a declining civic involvement. Research in service of cultivation theory has produced the view that television in particular fosters a media-centered view that diverges from reality and supports an image of the world, and one's community, is a "mean and scary place," hardly safe for "bowling alone." The only media use variable consistently related to such values in the study reported here is newspaper readership, which is positively correlated with values representing civic culture, social network and social capital. Those who read the newspaper value political participation, being a good American, being involved in the community, having good neighbors and tolerance of other people. Reading the newspaper also is positively correlated with measures of community attachment, community activities, belonging to organizations and both metro and neighborhood quality of life assessments. However, going out to see films, using other print media, and frequent use of the Internet also are important. The strong pattern of positive relationships between community activities and use of new communication technologies is particularly impressive, suggesting that the newest medium—the Internet--acts more like newspapers in stimulating involvement in the community as spectators if nothing else. Unlike the newspaper, Internet use is not linked to organizational ties in the community.

Clearly, face-to-face communication is important for the social ties that are the fabric of social capital. In the study reported here, we see that those who have strong values representing social order, social solidarity and social networks also are open to establishing relationships with strangers and are more likely to talk about current events and items raised by the media. Interpersonal communication variables also are related to almost all community variables, supporting strong community attachments, involvement in more organizations and positive assessments of the community quality of life. This pattern also fits the picture of a stable neighborhood painted by Doolittle and MacDonald (1978), who found a relationship between communication and sense of community.

Communication variables often are viewed by those outside the discipline as merely mediating influences from above or below. Those framing issues at the societal, or community level often position media and other communication variables in models where they largely reflect social categories or larger institutional and social arrangements. At the same time, some observers believe that audiences engage in selective exposure and project their values onto the media. Here we entered social categories, civic and community values, and communication variables in hierarchical regression to see to what extent media and interpersonal communication patterns had an impact on social capital beyond that explained by values and demographics. Regression analyses predicting a summary measure of social capital show that communication variables explain additional variance beyond that accounted for by social categories, civic values and community values. Among the communication variables, talking about things in the media and interpersonal relations—openness to strangers—are the most important predictors. Thus, both mass and interpersonal communication are important for building social capital.

Table 1
Correlations between Civic & Community Values and Community Variables

	Community attachment	Community activities	Org. ties	Metro QOL	Neigh QOL
Civic Culture Values					
Participating in political system	.21***	.13*	.25***	.21***	.13*
	.20***	.13*	.24***	.20***	.12*
Being a good American	.17**	-.09	.02	.10#	.19***
	.13*	-.03	.04	.09	.16**
Social Order & Control Values					
Tolerance of other people	.08	.03	.06	.21***	.21***
	.03	.06	.05	.17**	.16**
Being obedient to authority	.12*	-.11#	-.01	-.00	.12*
	.08	-.05	-.00	-.01	.10#
Social Solidarity Values					
Equality for all	.12*	-.09	-.06	.17**	.13*
	.09	-.07	-.07	.17**	.13*
Being able to help others	.18***	-.07	-.02	.02	.07
	.17***	-.05	-.03	.04	.08
Social Network & Social Capital Values					
Being involved in the community	.40***	.14*	.24***	.18**	.06
	.42***	.16**	.24***	.18**	.08
Having good neighbors	.27***	-.03	.11#	.15*	.22***
	.24***	.01	.11#	.13*	.20***

Note: The first line of figures represents bivariate correlations between the values and community variables, based on a sample size of about 300. The second line represents the partial correlations between values and community variables controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 2
Correlations between Civic & Community Values and Traditional Media Use Variables

	TV hours	TV news	Radio hours	Read paper	Read mags	Read books	Watch videos	See films	Media Index
Civic Culture Values									
Participating in political system	.05	.08	-.02	.27***	.08	.09	.05	-.03	.15*
	.06	.09	-.00	.25***	.07	.07	.08	-.02	.14*
Being a good American	.11#	.11#	-.04	.14*	-.07	.07	-.09	-.16**	-.02
	.08	.05	-.04	.12*	-.05	.05	.00	-.09	.01
Social Order & Control Values									
Tolerance of other people	.06	.05	-.06	.18**	-.01	.00	-.03	-.00	.04
	.07	.04	-.04	.12*	-.03	-.01	.03	.05	.06
Being obedient to authority	.06	.08	.03	.08	-.05	-.03	-.04	-.02	.01
	.03	.04	.03	.04	-.03	-.06	.05	.05	.04
Social Solidarity Values									
Equality for all	-.06	.01	-.04	.02	-.05	.04	-.02	-.10#	-.08
	-.06	.00	-.04	-.02	-.06	.03	.02	-.08	-.08
Being able to help others	-.02	-.01	.03	.08	-.10#	.02	-.03	-.15**	-.06
	-.04	-.04	.03	.06	-.11#	-.02	.01	-.11#	-.06
Social Network & Social Capital Values									
Being involved in the community	.05	.02	.10#	.20***	.06	.13*	.03	-.00	.17**
	.05	.02	.11#	.23***	.06	.13*	.05	.01	.19**
Having good neighbors	-.01	.06	.03	.12*	-.08	.01	-.07	-.14*	-.04
	-.03	.01	.04	.08	-.08	-.01	-.00	-.11#	-.03

Note: The first line of figures represents bivariate correlations between the values and traditional media use variables, based on a sample size of about 300. The second line represents the partial correlations between values and traditional media use variables controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 3
Correlations between Traditional Media Use Variables and Community Variables

	Community attachment	Community activities	Org. ties	Metro QOL	Neigh QOL
TV hours	.03	-.09	-.01	-.08	-.01
	-.00	-.07	-.02	-.08	-.01
TV news	**.10#**	-.03	.06	-.04	.09
	.06	.01	.03	-.04	.08
Radio hours	**.13***	.09	.02	-.00	.02
	.15*	**.10#**	.03	.02	.05
Read newspaper	**.23*****	**.21*****	**.23*****	**.12***	**.13***
	.18**	**.25*****	**.21*****	**.11#**	.06
Read magazines	.07	**.25*****	**.17*****	.09	.08
	.08	**.24*****	**.15***	**.10#**	**.10#**
Read books	.04	**.12***	**.15****	.02	**.13***
	.02	**.12***	**.13***	.05	**.15***
Watch videos	-.09	.08	.00	.04	-.08
	-.01	.01	.02	.07	-.02
Go out to see films	**-.14***	**.24*****	.06	-.00	-.08
	-.11#	**.20*****	.06	.01	-.05
Media Use Index	.08	**.28*****	**.19*****	.04	.04
	.09	**.26*****	**.17****	.06	.08

Note: The first line of figures represents bivariate correlations between the values and traditional media use variables, based on a sample size of about 300. The second line represents the partial correlations between values and traditional media use variables controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 4
Correlations between Community Variables and Use of New Communication Technologies

	Community attachment	**Community activities**	**Org. ties**	**Metro QOL**	**Neigh QOL**
Has computer	.08	.19***	.11#	.10#	.08
	.11#	.17**	.10	.12*	.07
Has gone on Internet before	-.12*	.19***	.00	.01	-.02
	-.05	.14*	.01	.03	.03
Has Internet access	.01	.22***	.08	.03	.01
	.08	.17**	.08	.06	.07
Internet use at work	-.03	.22***	.02	.07	.16**
	.03	.15*	-.00	.06	.19**
Internet use at home	.07	.22***	.03	.08	.14*
	.12*	.19***	.03	.05	.11#
Overall Internet use	.02	.25***	.03	.08	.17***
	.09	.20***	.01	.06	.18***
Chatroom use	-.08	.29***	.03	.02	-.03
	.00	.26***	.05	.02	.01
Freq. visits media websites	-.02	.19***	.06	.13*	.07
	.03	.14*	.05	.11#	.07

Note: The first line of figures represents bivariate correlations between the values and traditional media use variables, based on a sample size of about 300. The second line represents the partial correlations between values and traditional media use variables controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 5
Correlations between Community & Civic Values and Use of New Communication Technologies

	Has computer	Has gone on net	Has net access	Net use at work	Net use at home	Overall Net use	Chat room visits	Visits media websites
Civic Culture Values								
Participating in political system	.03	-.03	-.02	.03	.09	.07	.07	.00
	.01	-.05	-.04	.01	.07	.04	.09	-.01
Being a good American	-.09	-.25***	-.14*	-.23***	-.10#	-.19***	-.12*	-.20***
	-.04	-.16**	-.05	-.14*	-.03	-.10#	-.03	-.12*
Social Order & Control Values								
Tolerance of other people	.17**	-.06	.04	.05	.12*	.10#	-.06	-.03
	.17**	-.05	.06	.05	.10#	.09	-.01	-.03
Being obedient to authority	-.04	-.18**	-.07	-.21***	-.10#	-.18**	-.04	-.13*
	-.02	-.11#	-.01	-.14*	-.05	-.11#	.05	-.05
Social Solidarity Values								
Equality for all	.04	-.04	.02	-.05	.02	-.02	-.10#	.01
	.02	-.06	.01	-.07	-.00	-.04	-.10#	.01
Being able to help others	-.09	-.13*	-.10#	-.12*	-.10#	-.13*	-.20***	-.11#
	-.09	-.11#	-.09	-.08	-.08	-.10	-.16**	-.06
Social Network & Social Capital Values								
Being involved in the community	.08	-.01	.03	.04	.05	.05	.06	.07
	.08	-.05	.02	.03	.03	.04	.04	.06
Having good neighbors	.02	-.12*	-.07	-.14*	.01	-.08	-.12*	-.11#
	.04	-.05	-.02	-.10#	.05	-.03	-.06	-.06

Note: The first row represents bivariate correlations between use of new communication technologies and civic and community values, based on a sample size varying slightly from 300. The second row represents partial correlations between use of new communication technologies and civic and community values while controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 6
Correlations between Interpersonal Communication Variables, and Community & Civic Values, and Community Variables

	Interpersonal Relationships		Interpersonal Topics	
	Enjoy meeting Strangers	Prefer talking to people like self	Talk freq. about current events	Talk freq. about things in media
Civic & Community Values:				
Civic Culture Values				
Participating in political system	.08	.05	.05	.05
	.06	.05	.03	.04
Being a good American	.06	.08	.06	.05
	.02	.08	.09	.10
Social Order & Control Values				
Tolerance of other people	**.25*** **	-.02	**.23*** **	**.14***
	.20* **	-.05	**.21*** **	**.15***
Being obedient to authority	**.15***	.06	**-.13***	-.06
	.12*	.08	**-.11#**	-.03
Social Solidarity Values				
Equality for all	**.25*** **	-.08	**.12***	**.14***
	.24* **	-.05	**.12***	**.14***
Being able to help others	**.11***	-.02	.01	**.14***
	.11#	-.00	.02	**.16***
Social Network & Social Capital Values				
Being involved in the community	**.19*** **	-.04	**.14***	**.21*** **
	.20* **	.01	**.13***	**.20*** **
Having good neighbors	**.21*** **	**.12***	**.14***	**.12***
	.18 **	**.10#**	**.14***	**.13***
Community Variables:				
Community attachment	**.24*** **	**.19** **	**.15***	**.12***
	.20* **	**.16** **	**.15***	**.24*** **
Community activities	.07	.09	**.21*** **	**.24*** **
	.12*	.09	**.20*** **	**.24*** **
Organizational ties	**.16** **	.08	**.15***	**.18** **
	.17 **	.06	**.13***	**.17** **
Metro QOL assessment	**.31*** **	.09	**.20*** **	.09
	.28* **	.04	**.17** **	.08
Neighborhood QOL assessment	**.17** **	**.18** **	**.17** **	**.10#**
	.11#	**.14***	**.14***	**.12***

Note: The first row represents bivariate correlations between interpersonal communication variables, and civic and community values and community variables, based on a sample size varying slightly from 300. The second row represents partial correlations between interpersonal communication variables, and civic and community values and community variables, while controlling for social categories, including: gender, ethnicity (white, black, other ethnic status as dummy variables), life cycle (age, married marital status, number of people in household), and achievement (education and income); the sample size varies slightly from 270. #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

Table 7
Predicting Social Capital

	R^2 Ch.	F Ch.
Dependent Variable= Factor score for Social Capital[1]		
Social Categories	.052	1.5 n.s.
Civic & Community Values	.169	6.3 p<.001
Communication Variables	.153	3.5 p<.001
	Equation: R=.61, R^2=.37, F=4.1, p<.001, N=250.	
Dependent Variable=Organizational Ties[2]		
Social Categories	.041	1.2 n.s.
Civic & Community Values	.105	3.6 p<.001
Communication Variables	.071	1.3 n.s.
	Equation: R=.47, R^2=.22, F=1.9, p<.004, N=250.	
Dependent Variable=Community Attachment[3]		
Social Categories	.065	1.9 p<.06
Civic & Community Values	.185	7.2 p<.001
Communication Variables	.088	1.9 p<.02
	Equation: R=.58, R^2=.34, F=3.5, p<.001, N=250.	
Dependent Variable=Community Activities[4]		
Social Categories	.084	2.5 p<.01
Civic & Community Values	.048	1.5 n.s.
Communication Variables	.165	3.4 p<.001
	Equation: R=.54, R^2=.30, F=2.9, p<.001, N=250.	
Dependent Variable=Quality-of-Life Assessments[5]		
Social Categories	.375	4.4 p<.001
Civic & Community Values	.096	3.7 p<.001
Communication Variables	.081	1.7 p<.05
	Equation: R=.56, R^2=.32, F=3.2, p<.001, N=250.	

Note: The table includes five stepwise regressions where three sets of predictor variables were entered in the sequence listed above, with social categories first, civic and community values second and communication variables third. For the standardized betas listed below, #=p<.10; *=p<.05; **=p<.01; ***=p<.001.

[1] The summary measure for social capital is the factor score resulting from a factor analysis of organizational ties, community attachment, community activities and QOL assessments, where one factor emerged. Statistically significant standardized betas by step were: Social Categories, none; Civic & Community Values, participating in the political system, .17**, being involved in the community, .30***; having good neighbors, .12#; Communication Variables, frequency talk about things in media, .21**, enjoy meeting strangers, .17**, prefer talking with people of same background, .15*.

[2] Statistically significant standardized betas by step were: Social Categories, education, .18*; Civic & Community Values, participating in the political system, .18**, being involved in the community, .20**; Communication Variables, enjoy meeting strangers, .16*.

[3] Statistically significant standardized betas by step were: Social Categories, age, .23***; Civic & Community Values, being involved in the community, .36***; Communication Variables, radio listening, .12*, Internet use, .16#, frequency talk about things seen in media, .18*, prefer talking with people of same background, .13*.

[4] Statistically significant standardized betas by step were: Social Categories, education, .14*, being married, -.12#; Civic & Community Values, being involved in the community, .18*; Communication Variables, newspaper reading, .16*, going out to see films in theaters, .16*, visiting chatrooms, .17*, frequency talk about things seen in media, .20*.

[5] Statistically significant standardized betas by step were: Social Categories, household income, .12#, age, .14*, white ethnicity, .53***, African-American ethnicity, .30*, other ethnicity, .27*; Civic & Community Values, participating in the political system, .15*, equality for all, .14#, having good neighbors, .13#; Communication Variables, frequency talk about current events, .16#, enjoy meeting strangers, .16*, prefer talking with people of same background, .13*.

References

Althaus, S.L., & Tewksbury, D. (2000). Patterns of Internet and traditional news media use in a networked community. Political Communication, 17, 21-45.

Andrews, F.M. (Ed.) (1986). Research on the quality of life. Ann Arbor: University of Michigan, Institute for Social Research.

Ball-Rokeach, S.J., Kim, Y., & Matei, S. (2001). Storytelling neighborhood: Paths to belonging in diverse urban environments. Communication Research, 28, 392-428.

Bandura, A. (2001). Social cognitive theory of mass communication. Media Psychology, 3, 265-299.

Becker, B., & Wehner, J. (1998). Electronic media and civil society. Electronic Journal of Communication, 8, ¾.

Brintnall, M., Folkerts, J., & Musson, T. (1990). News about state community block grant programs: A survey of state directors. Journalism & Mass Communication Quarterly, 67, 163-170.

Demers, D.K. & Viswanath, K. (Eds.) (1999). Mass media, social control and social change: A macrosocial perspective. Ames: Iowa State University Press.

Docherty, I., Goodlad, R., & Paddison, R. (2001). Civic culture, community and citizen participation in contrasting neighbourhoods. Urban Studies, 38, 2225-2250.

Doolittle, R.J., & MacDonald, D. (1978). Communication and a sense of community in a metropolitan neighborhood: A factor analytic examination. Communication Quarterly, 26, 2-7.

Emig, A.G. (1995). Community ties and dependence on media for public affairs. Journalism and Mass Communication Quarterly, 72, 2, 402-411.

Fessler, D.R. (1952). The development of a scale for measuring community solidarity. Rural Sociology, 17, 144-152.

Finnegan, J.R. Jr., & Viswanath, K. (1988). Community ties and use of cable TV and newspapers in a midwest suburb. Journalism Quarterly, 65, 456-473.

Fong, E., & Isajiw, W.W. (2000). Determinants of friendship choices in multiethnic society. Sociological Forum, 15, 249-271.

Forrest, R., & Kearns, A.(2001). Social cohesion, social capital and the neighborhood. Urban Studies, 38, 2125-2143.

Geiger, W., Bruning, J., & Harwood, J. (2001). Talk about TV: Television viewers' interpersonal communication about programming. Communication Reports, 14, 49-57.

Hyland, S.E., & Ciaramitaro, B. (1984). Developing new modes of communication in low-to-moderate income neighborhoods: Issues in energy conservation and folk beliefs. Journal of Voluntary Action Research, 13, 31-41.

Jeffres, L.W. (2002). Urban Communication Systems: Neighborhoods and the Search for Community. Cresskill, NJ: Hampton Press.

Jeffres, L.W., & Dobos, J. (1995). Separating people's satisfaction with life and public perceptions of the quality of life in the environment. Social Indicators Research, 34, 181-211.

Jeffres, L.W., & Dobos, J. (1984). Communication and neighborhood mobilization. Urban Affairs Quarterly, 20, 97-112.

Keyes, L. (1987-1988). The shifting focus on neighborhood groups: The Massachusetts experience. The Policy Studies Journal, 16, 300-306.

McCombs, M. (1997). Building consensus: The news media's agenda-setting roles. Political Communication, 14, 433-443.

McLeod, J.M., Daily, K., Guo, Z., Eveland Jr., W.P., Bayer, J., Yang, S. & Wang, H. (1996). Community integration, local media use and democratic processes. Communication Research, 23, 179-209.

Meegan, R., & Mitchell, A. (2001). 'It's not community round here, It's neighbourhood': Neighbourhood change and cohesion in urban regeneration policies. Urban Studies, 38, 2167-2194.

Miller, D.C. (1991). Handbook of research design and social measurement. 5th ed. Newbury Park, CA: Sage.

Moy, P., Pfau, M., & Kahlor, L. (1999). Media use and public confidence in democratic institutions. Journal of Broadcasting & Electronic Media, 43, 137-158.

Moy, P., Scheufele, D.A., & Holbert, R.L. (1999). Television use and social capital: Testing Putnam's time displacement hypothesis. Mass Communication & Society, 2, 27-45.

O'Hara, S.U. (2001). Urban development revisited: The role of neighborhood needs and local participation in urban revitalization. Review of Social Economy, 59, 23-43.

Orr, M., & West, D.M. (2002). Citizens' view on urban revitalization: The case of Providence, Rhode Island. Urban Affairs Review, 37, 397-419.

Peyrot, M., & Fenzel, M. (1994). Neighborhood surveys as community organizing. Paper presented to the American Sociological Association.

Pfau, M., Moy, P., Radler, B., & Bridgement, M.K. (1998). The influence of individual communication media on public confidence in democratic institutions. Southern Communication Journal, 63, 91-112.

Putnam, R.D. (1995a). Bowling alone: America's declining social capital. Journal of Democracy, 6, 34-48.

Putnam, R.D. (1995b). Tuning in, tuning out: The strange disappearances of social capital in America." PS: Political Science and Politics, 284, 664-683.

Putnam, R.D. (1996). The strange disappearance of civic America. The American Prospect, 24, 65-78.

Putnam, R.D., & Yonish, S. (1997). New Evidence on Trends in American Social Capital and Civic Engagement: Are We Really Bowling Alone?" Unpublished Manuscript, Harvard University. Cited in Shah et al. (1999).

Scheufele, D.A. (2002). Examining differential gains from mass media and their implications for participatory behavior. Communication Research, 29, 46-65.

Shah, D.V., Holbert, R.L., & Kwak, N. (1999). Expanding the "Virtuous Circle" of Social Capital: Civic Engagement, Contentment, and Interpersonal Trust." Paper presented at the annual conference of the Association for Education in Journalism and Mass Communication, New Orleans, LA.

Shah, D.V., McLeod, J.M., & Yoon (2001). Communication, context, and community: An exploration of print, broadcast and Internet influences. Communication Research, 28, 464-506.

Simonson, P. (1996). Dreams of democratic togetherness: Communication hope from Cooley to Katz. Critical Studies in Mass Communication, 13, 324-342.

Sotirovic, M., & McLeod, J.M. (2001). Values, communication behavior, and political participation. Political Communication, 18, 273-300.

Stamm, K.R. (1985). Newspaper use and community ties: Toward a dynamic theory. Norwood, NJ: Ablex.

Stamm, K.R., Emig, A.G., & Hesse, M.B. (1997). The contribution of local media to community involvement. Journalism & Mass Communication Quarterly, 74, 97-107.

Stocker, R., Green, D.G., & Newth, D. (2001, October). Consensus and cohesion in simulated social networks. Journal of Artificial Societies and Social Simulation, 4, 4.

Wilson, R.W. (1997). American political culture in comparative perspective. Political Psychology, 18, 483-502.

Endnotes

---

[1] A recent model simulation points out that critical levels of connectivity in neighborhood communication networks are required to achieve consensus, and neighborhood size is a factor (Stocker, Green & Newth, 2001).

[2] As Bandura (2001) notes, in social cognitive theory communication influences human thought, affect and action through two paths: 1) a direct path where communication informs, enables, motivates and guides people; and 2) a mediated path where participants are linked to social networks and community settings that provide natural incentives and continued personalized guidance.

[3] Also see the recent discussion about civic engagement and whether Americans' involvement in public activities has declined. Putnam (1995) argued in an essay that traditional civic engagement in America has been on a long, slow decline, but Stengel (1996) argues the reverse, that civic engagement isn't disappearing but reinventing itself.

[4] This was operationalized as attending club meetings, doing volunteer work, and participating in community projects.

[5] This included items tapping whether respondents felt people are honest and similar items for level of trust in government and business institutions.

[6] Shah, McLeod and Yoon (2001) found that social-recreational uses of the media were negatively related to several civic indicators while informational uses of the media were positively related.

[7] Simonson (1996) discusses the "dream" that mass communication will create a national community by media attention to American ideals of community, communication and democracy.

[8] Their research identifies five dimensions: psychological attachment (like living in area, view it as home, likelihood will move away), interpersonal discussion networks (indexed by discussion with neighbors, discussing area problems, getting together with neighbors and the proportion of friends who live in the area), city versus group (identify with city or with social group), localism versus cosmopolitanism (local news more interesting than national, best organizations are local), and city versus neighborhood (concern with larger community rather than concern with local town or neighborhood and identification with neighbors).

[9] Sotirovic and McLeod (2001) found that reflecting about news and integrating the information gained with that from other sources promotes better understanding of the world of politics and may provide the basis for political participation that is stronger than that achieved through mere factual knowledge gain.

[10] Several of these domains include community-based values similar to the features of political culture identified by Wilson (1997).

---

[11] A random sample of phone numbers was drawn from the metropolitan telephone book. Then the last two digits were randomly assigned using a random numbers table.

CT&M

None of the Above:

Creating Mass Deliberation Without Discussion

Ray Pingree
Graduate student
School of Journalism and Mass Communication
University of Wisconsin-Madison
ray_pingree@hotmail.com
(608) 244-3336
210 Buell St #2
Madison WI 53704

None of the Above:

Creating Mass Deliberation Without Discussion

Abstract

Deliberative democracy has been plagued by questions of implementation, due to a failure to distinguish between discussion and the more general concept of many-to-many communication. To demonstrate that this theoretical distinction is both possible and important, this paper introduces an example of an Internet-based many-to-many communication system designed to achieve deliberation's outcomes without discussion. A broader deliberative theory is proposed, to encompass the concept of non-conversational deliberation as part of a more attainable public sphere.

"[Online forums are] a development of historic significance, for there has been practically no innovation in many-to-many communication in over two thousand years"
- Hans Klein (1999: 213)

"The main problem [for deliberative democracy] is to explain how it is even possible to have a 'discussion' among thousands or millions of people."
- James Fearson (1998:64)

None of the Above:

Creating Mass Deliberation Without Discussion

Deliberation and discussion

The promise of deliberative democracy is in its ideal outcomes, but its most compelling criticisms highlight the unreliability of discussion for achieving those outcomes, especially in large, diverse groups or an entire polity (Witschge 2002, Shapiro 1999, Bell 1999, Hardin 1999, Sanders 1997). Similarly, both optimists and pessimists about the deliberative potential of the Internet have focused on the strengths and weaknesses of online discussion (Papacharissi 2002, Witschge 2002, Dahlberg 2001, Fishkin 2000, Gastil 2000, Davis 1999, Klein 1999). Underlying much of the disagreement about the deliberative potential of the Internet is a disagreement about the deliberative potential of discussion itself.

By discussion I mean any group communication structured as a series of conversational turns. Because of the Internet, discussion is no longer the only way to structure group communication, although all previously existing forms of online group communication are discussion-based (including email, Usenet, chat, message boards, and blogs). Entirely new forms of online group communication can be designed that directly address the goals of deliberation. To persuade the reader on this theoretical point requires a concrete example, and since no such example already exists, this paper will introduce a new software system designed to be such an example.

In deliberative democratic theory, "deliberative" does not refer to just any discussion. Different theorists list different conditions that a discussion has to satisfy to be called deliberative (Witschge 2002, Elster 1998, Fearson 1998). What is important for our purposes is not these conditions themselves, but the fact that arguments over their appropriateness are nearly always based on their capacity to encourage smarter or fairer collective decision-making. This points to a more general, although tautological, definition of deliberative: a group communication process is deliberative to the extent that it transforms (not merely aggregates) individuals in such a way as to make society, as a whole, behave better in terms of collective intelligence and justice.

To create a broader deliberative theory, one that is not biased towards discussion, we must work backwards from ideal outcomes of communication to discover ideal communication conditions. The first backwards step is therefore a clearer definition of the goals of political communication than "an intelligent and just society." To that end, I suggest two intermediate outcomes that political communication processes should aim for: collective reasoning and pluralistic awareness.

Collective reasoning

Collective reasoning, at its heart, is the sharing of statements and considerations that bear on them. We borrow the spirit of Zaller's concept of a consideration: basically anything that bears on a decision to agree or disagree with a statement (1992:40-41). Zaller's concept, however, was intended for how public opinion is usually measured: fixed survey questions (1992:4), and thus quite rightly does not take into account a third reaction (beyond merely agreeing or disagreeing) that people often have when they encounter an opinion-statement: the desire to re-frame or re-phrase it.

Note that as group size grows, discussion must tend more towards Zaller's definition of considerations. Nonverbal cues can express agreement or disagreement, but one must take up a conversational turn to re-frame. With a bigger group, each individual has a smaller fair share of speaking time, so each individual gets fewer opportunities to re-frame as group size grows. In the extreme case of a national (or global) "discussion," mass opinion can only be agreement or disagreement with elite-framed opinion (which includes both opinion polling and voting). These are limits that follow directly from the structure of conversation, and should not be presupposed to be limits of group communication in general. For this reason, our definition of consideration must include not only reasons to agree and reasons to disagree, but also alternative frames. Considerations then bear not only on the decision of whether to agree or disagree, but also on decisions such as whether to even answer the question as asked, and what alternative questions to propose.

Reasoning as exploration

Each consideration is potentially a whole new opinion-statement, which may in turn have its own long list of possible considerations. Thus, reasoning (alone or with others) can be thought of as exploring a network. At any point in this exploration, we may not be aware of all the possible pathways (or considerations) we could take. As Zaller argues, the sets of considerations that occur to people when asked to privately evaluate an opinion statement tend to be unstable over time and vulnerable to salience effects like question order, question wording, and media priming (1992:40-96). These limits of private reason constitute a strong argument for deliberation: by pooling considerations, we can form better opinions. Instead of wandering this vast network of opinions and considerations alone, we attempt to bring a group along with us, so that at every step we have the benefit of others' vision (and experience) about which way to go.

In this exploration, discussion requires a group to generally "stick together" in order to benefit from each other. This is why conversational norms, civic practices (Eliasoph 1996), and discussion rules are crucial to conversational deliberation. This is also why discussion doesn't scale. With the Internet, we can design software that for the first time allows explorers to benefit from anyone who has ever been "in the same place" before, while allowing them to individually "go wherever they want" at any time. In other words, by structuring a system as a network of opinion statements connected by considerations, instead of as a series of conversational turns, we can allow people viewing a particular opinion statement to benefit from the reasoning of anyone who has ever viewed that statement in the past, while allowing each individual to view or write whatever opinions or considerations they want to, at any time. If we don't get along well with each other in such a system, it does not have to impede our ability to benefit from each other's reasoning.

Note that existing message boards and blogs are asynchronous and have links, but these links are still generally to and from conversational turns. Users can "go where they want" independently of each other, but in a conversation network, not a consideration network. In other words, instead of seeing considerations that prior visitors reading the same thing have thought were relevant to deciding what to make of the statement, users see a history of what other statements have been made in response. Reasons to agree or disagree with a statement may be buried deep in the many replies and sub-replies to the statement, making it difficult to weigh them against each other. Furthermore, such systems do nothing to help people see which reasons have been seen as compelling by others.

Collective reasoning creates the public

A key difference between collective reasoning and private reasoning concerns the public-spiritedness of the reasons themselves, via what Elster calls "the civilizing force of hypocrisy" (1998:12). As dissatisfied as I am by rational choice theory in general, it's hard to deny that some people sometimes privately reason solely in self-interest. However, when addressing what is perceived to be a diverse audience, even a purely selfish person has an incentive (in order to be persuasive) to express public-spirited reasons (reasons based on some concept of the common good). Through collective reasoning, we "form in common a common will" (Elster 1998:2). Public-spirited reasoning, even if it does not broaden consensus, forces us to create some version of a public spirit in each of our individual heads, resulting in a real difference in collective understanding and political coherence. This is what Eliasoph refers to as "the power to create the public itself" (1996:263).

The question of whether or not people are rational actors is analogous to the question of what peoples' "true opinions" are. In other words, Zaller's model stressing the importance of the salience of considerations (1992) applies to selfishness too. If, on balance, selfish considerations are more salient than public-spirited ones at the moment, an individual is more likely to make a self-interested decision. This would suggest that even if people currently seem to behave mostly self-interestedly, this may not be so much an indicator of human nature, but simply a lack of salience (or even lack of knowledge) of public-spirited considerations, due to living in a nearly deliberation-free society. Therefore, there is some hope that if we create a public, public-spiritedness (and thus further re-creation of the public) will be easier.

Pluralistic awareness

> "People need an organized map of the political world, not just a huge pile of unsorted facts" (Eliasoph 1998:152).

By pluralistic awareness, I mean knowledge of the structure of opinions in society. Through communication we can learn this structure. This includes not only how many people agree or disagree (and how passionately) with each opinion statement or consideration, but also which sets of opinions tend to go together in opinion-groups (e.g. the left, the right), which opinions bridge those opinion-groups, and which opinions are marginal.

In other words, we can form, through communication, a mental map of opinion space. However, in *discussion*, it may also go the other way. The mental map may affect our conversational turns and our interpretations of the turns of others. Through the opinion-structural equivalent of Noelle-Neumann's Spiral of Silence (1984), we may reproduce our own assumptions about what "doesn't fit" into assumed political categories, and what would thus be difficult to explain in a limited conversational turn. What we say, how we say it, and how we hear what is said can all depend on (and determine) "where" we perceive other discussants to be in our mental map.

For example, the common assumption that opinion space consists of a left, a right, and a center may lead people to converse more as if they are addressing and representing those categories than they would otherwise, especially as group size grows. People would do this to save time, because of the aforementioned limits of the conversational turn. Even if nobody "misrepresents" any of "their own" reasons or opinions, in conversation one is forced to choose which of many possible opinions and considerations to express in each conversational turn. Again, the key point is that this can socially reproduce and reinforce that mental map of left, right, and center (or whatever the current map is), regardless of what would happen to the map given an open exchange.

The effects on a democracy of having an incorrect yet persistent shared mental map would be enormous. Nearly everyone would feel like they personally "didn't fit" into politics, so participation would be low, as would trust in established political parties and government. Attempts at collective reasoning would be unsatisfying and might become either taboo or polarized to the point where they resemble a sporting event, only with less sportsmanship.

Interactions of collective reasoning and pluralistic awareness

The combination of collective reasoning and pluralistic awareness is critical. A society composed of two groups that each misunderstands the other is clearly worse off than one that is equally polarized but where both sides know the reasons why. Thus, the sharing of considerations is not just to help people decide where they stand, but to help people to understand why other groups of people stand where they do.

Deliberation should not be expected to lead to consensus, because of deep differences that can't be resolved by reasoning, and thus will still exist between "reasonable" people. Such differences include values, vested interests, tastes, and religious beliefs. Cohen calls this "the fact of reasonable pluralism" (1997:408), and argues that it is here to stay (1993, 1997, 1998). But even without consensus, deliberation should be expected to lead to increased understanding of such deeper reasons for different opinions, resulting in increased political tolerance and higher chances of finding good compromises between truly meaningful groups (in other words, groups formed out of shared awareness deep reasons), based on a mutual understanding of the reasons for their differences. Thus, deliberation must combine collective reasoning and pluralistic awareness to lead to political tolerance. This may explain part of the unreliability of effects of conversational deliberation on tolerance (for example, compare Denver et al. 1995 to Fishkin et al. 1999).

Both collective reasoning and pluralistic awareness make use of spatial metaphor. It's important to note that they refer to two different, but related "spaces," or more precisely, networks. In collective reasoning, people explore together a *semantic* network. It is made up of opinion statements connected by considerations. Pluralistic awareness is awareness of a different space. It is ultimately based on assumptions of *statistical* relationships between pairs of opinion statements (e.g. people who agree with A also tend to agree with B, or people who agree with C tend to disagree with D). Our mental maps of opinion structure may take various forms (category systems, dimensions, or more likely some fuzzy combination of the two), but I argue that the underlying data on which we would ideally base those mental maps is actually the statistical relationships between pairs of individual opinion statements.

The Meaning Map Project

Pessimists about the deliberative potential of the Internet have often focused on what Internet software currently does or how it is currently used (e.g. Davis 1999). Imagine doing the same thing in the early days of electricity. Asking what effects electricity will have might have seemed to make sense when the only thing electricity could do was light a light bulb. Today we know electricity can be used for many different things, so to have predicted the effects of all devices that have used electricity would have required some technological vision about what those devices might be. The same is true of the Internet. To evaluate the Internet's potential requires technological vision about what kinds of software could be developed using the Internet. It is not sufficient to analyze the software that currently uses it, the ways that that software is currently used, or the gratifications sought by its users. To assess the potential of the Internet to achieve specific goals, we must attempt to design Internet software specifically for those goals, and then attempt to test whether those goals are achieved in actual use.

I have tried to do exactly that for the goals of facilitating collective reasoning and pluralistic awareness in large groups. The resulting software is currently in alpha testing, and will soon be available for public use and experimentation at www.meaningmap.com.

The same website will also soon have more detailed descriptions of the system. I give only a brief overview of the system and some of its limitations here, in order to demonstrate the feasibility of non-conversational deliberation and to inspire other social scientists to try their hands at software design.

Overview

The Meaning Map is an online open polling system, where users explore opinion statements either via a visual map of the network of statistical relationships between opinions, or by traversing individual consideration-links in a collaboratively created semantic opinion network. Unlike an opinion poll with fixed questions, anyone can at any time post a new opinion statement, which other users can then cast agree or disagree votes on. The patterns of these votes solely determine the statistical opinion network. Instead of being forced to agree or disagree with each opinion, users can also link it to another opinion that they feel is a better way of framing the issue. These "better frame" considerations appear in a list visible to all users viewing that opinion in the future. In another list are the other two kinds of considerations: reasons to agree and reasons to disagree. Users can vote on the quality of each consideration, determining the order of display of the consideration list (but not affecting the opinion's position in the map).

Users are free to navigate on their own, and to post new opinion statements or considerations regardless of what any other user is doing. When viewing any opinion statement, they have the benefit of a shared and rank-ordered list of the considerations added by other users who have viewed that statement in the past. Reasoning together in this way makes getting along with others far less important and ensures that the results of the collective reasoning effort are preserved in a meaningful structure, instead of being scattered in a disorganized history of replies.

Since statistical relationships in the patterns of votes on opinion statements are visually displayed in a map for all to see, users of this system should have a more accurate and up-to-date form of pluralistic awareness than one can glean from anecdotal evidence of the known opinion-sets of discussion partners.

Visualizing the statistical network

The technical challenge here is how to communicate (through both display and interaction) to ordinary people that opinion "space" is best thought of as a network of pair-wise relationships between individual opinions. It should not be presupposed that it is based on a handful of latent factors (dimensions) or latent classes (categories). Such categories or factors may be deduced from exploration of the network, but a network structure cannot be deduced from exploration of categorized or factored data.

In network visualization, the goal is usually to create a still image where the distances between pairs of points in the image correspond as closely as possible to their network distances. In our system, we do not need to limit ourselves to a still image. Through animation and user interaction, the process by which points find their positions (repulsion from negatively related opinions and attraction towards positively related opinions) is revealed to the user. Users can drag and drop individual points around the screen, allowing them to see other places where an opinion might come to rest in the display (and see indirect effects on the rest of the space due to that change). Also, users can temporarily show only a subset of opinions, which may result in a dramatic re-ordering of their positions when outside influences are removed. This allows the user to see the internal structure of some meaningful set without reference to other opinions (for example, seeing that socialism is central to the left when the left is viewed on its own, but peripheral in a broader view, because it is strongly repelled by the right).

Filtering and statistical zoom

> "Any form of filter imposes its own biases. But the absence of any filter also has its own bias. It causes public opinion expression to break down into a babel of voices, with only the loudest achieving some level of recognition" (Davis 1999:166).

Not all opinions are displayed in the opinion space at once, unless the topic is very new. This is not so much due to computational constraints as to the fact that unless the network structure is "simple," beyond a certain number of opinions, their positions become so stressed (pulled and pushed in so many directions) as to be meaningless. By "simple" I mean driven almost entirely by a handful of underlying categories or dimensions (instead of the complex webs of semantic relationships implicit in the idea of considerations). If the opinion space actually is dimensionally simple (for example a left-right continuum or the left-right / authoritarian-libertarian "Nolan chart"), network analysis methods can uncover these dimensions more accurately than factor analysis (Brazill and Grofman 2002).

In case the space is not simple, we need tools to explore it as a network. First, this means displaying a reasonable number of points at once. The Meaning Map displays the most prominent opinions as the "top level" view. The system allows the user to specify by what combination of criteria each opinion is rated as prominent.

For the same reasons, the system also indicates node-level stress. The color of opinions indicates the extent to which they "fit" in their current position. Opinions that fit well are blue, and as stress increases the color fades to purple and then yellow. This highlights opinions that serve as bridges between clusters or as "wormholes" in the dominant dimensions.

Finally, the system allows for exploration of the network by providing a "statistical zoom." Since underlying dimensions cannot be assumed from the outset, the idea of "zooming in" to view less prominent opinions "near" a certain opinion can only be based on their statistical nearness. When a user selects an opinion to view in detail (in order to read considerations and/or give an informed vote), they see a map of a different set of opinions, selected not just for prominence but also for their statistical relatedness to the current opinion.

Statistical zoom can result in a dramatic re-ordering of the space – for example, "god exists" may be a highly stressed bridging opinion between left and right, both negatively and positively related to many opinions in those two dominant clusters. However, when this opinion is made artificially dominant by statistically zooming to it, one might see a completely different ordering, where religious and non-religious (or moralistic and non-moralistic) statements are two dominantly opposed clusters. This can be a powerful (if indirect) way of seeing different worldviews in the same data.

Limitations

The Meaning Map is a relatively simple first step. The set of problems in deliberative democracy it attempts to address do not include those of constructing a full alternative political system. Apart from issues of translating the results of national deliberation into national policy, many technical issues would need to be addressed before this system could even be capable of creating one integrated national deliberation. And then there are security, identity and privacy issues. This system merely elevates the possible deliberative group size from on the order of tens, to perhaps tens of thousands, and makes deliberative outcomes more likely among diverse groups of strangers. Even this simple system would be very useful to citizens trying to pool their abilities to evaluate existing or proposed legislation. It would also be useful to visually see candidates' positions (and their change over time) on the most prominent issues, where that agenda is determined by a deliberation among citizens instead of being driven by candidate strategy or by media agenda setting.

The Meaning Map does not attempt to directly facilitate empirical verification or the evaluation of risks (automating the combination of probabilities and severities), although it clearly can help large groups of people do both of these things more efficiently than discussion. Some potentially important work on these possibilities is taking place on the topic of deliberation for environmental risk regulation. Payne (1996) points out the suitability of deliberation for global environmental issues, and McBurney and Parsons (2001) formulate an environmental-risk-specific ideal deliberative system, which, although conversational in structure, represents a compelling argument for customized public sphere facilitation systems for specific domains.

At the time of this writing, the opinion mapping methods have been tested with real data, but the system has not yet been tested with a large group of users. Unforeseen problems of all varieties could arise through real use. It is conceivable that exposure to the opinion map may lead to an increase instead of a decrease in polarization. Experiments should be conducted to assess the impact of four different experimental conditions: the complete system, the system without the statistical map, the system without considerations, and the system without either (just a simple open polling system). Collaboration on these or other experiments with this software is very welcome.

Conclusion

Richard Davis argues that "the notion that the public will take control of agenda setting is absurd. … Someone must *organize the discussion* and *frame the alternatives*. Then, and only then, can the public respond intelligently" (1999:170, emphasis added). If we can complete the work this project begins, making mass deliberation technically feasible, the situation will be reversed. It will instead be absurd to claim that the public should not participate in setting its own agenda, that aggregates of privately-measured agreements with fixed opinion statements framed by elites can be called *public* opinion, or that elections conducted under such conditions can be called legitimate democracy.

Which sphere is virtual?

Recent work on the deliberative potential of the Internet has used the term "virtual sphere" to refer to online discussion (Papacharissi, 2002). Zizi Papacharissi's overview of the subject highlights the differences between the Habermasian ideal public sphere and the realities of discussion, both online and offline. She summarizes Michael Schudson's 1997 critique of deliberation-as-discussion: "there is little evidence that a true ideal public ever existed, and … public discourse is not the soul of democracy, for it is seldom egalitarian, may be too large and amorphous, is rarely civil, and ultimately offers no magical solution to problems of democracy" (Papacharissi 2002:11). It's high time for us optimists to admit that discussion itself may not be an adequate tool for creating the ideal public sphere.

However, our goal can no longer be seen as a "tragic and stoic pursuit of an almost impossible rationality, recognizing the impossibility of an ideal public sphere and the limits of human civilization, but still striving toward it" (Papacharissi 2002:11). More important than whether communication occurs online or offline are the questions of whether communication is guided by arbitrary, oversimplified, and self-reproducing mental maps of opinion space, whether people must reason in lock-step together in order to benefit from (and come to understand) each other's reasoning, and whether everyone can participate equally in the framing and agenda setting processes. The possibility of removing these limits creates a whole new set of opportunities and problems, both theoretical and technical, which may lead to a more real public sphere than has ever existed before.

References

Bell, D. (1999). Democratic Deliberation: The Problem of Implementation. In Macedo, S. (ed.), *Deliberative Politics: Essays on Democracy and Disagreement*, 70-87.

Brazill, T. and Grofman, B. (2002). "Factor analysis versus multi-dimensional scaling: binary choice roll-call voting and the U.S. Supreme Court." *Social Networks* 24:201-229.

Cohen, J. (1993). Moral Pluralism and Political Consensus. In D. Copp, J. Hampton, and J. Roemer (eds.), *The Idea of Democracy.* Cambridge University Press, 270-91.

(1997). Procedure and Substance in Deliberative Democracy. In J. Bohman and W. Rehg (eds.), *Deliberative Democracy: Essays on Reason and Politics.* Cambridge, MA: MIT Press

(1998). Democracy and Liberty. In J. Elster (ed.), *Deliberative Democracy.* Cambridge University Press, 185-231.

Dahlberg, L. (2001). Computer-Mediated Communication and The Public Sphere: A Critical Analysis. *Journal of Computer-Mediated Communication*, 7(1), 1-28.

Davis, R. (1999). *The Web of Politics.* New York: Oxford University Press.

Denver, D., Hands, G., & Jones, B. (1995). Fishkin and the deliberative opinion poll: Lessons from a study of the Granada 500 television program. *Political Communication*, 12, 147-156.

Eliasoph, N. (1996). Making a Fragile Public: A Talk-Centered Study of Citizenship and Power. *Sociological Theory*, 14, 262-289.

(1998). *Avoiding Politics: How Americans Produce Apathy in Everyday Life.* New York, NY: Cambridge University Press.

Elster, J. (1998). Introduction. In J. Elster (ed.), *Deliberative Democracy.* Cambridge University Press, 1-18

Fearson, J. (1998). Deliberation as Discussion. In J. Elster (ed.), *Deliberative Democracy.* Cambridge University Press, 44-68

Fishkin, J., & Luskin, R. (1999). Bringing deliberation to the democratic dialogue. In M.

McCombs, A. Reynolds, & Eds. (Eds.), *The poll with a human face: The National Issues Convention experiment in political communication,* Mahwah, NJ: Lawrence Erlbaum, 3-38

Fishkin, J. (2000). Virtual Democratic Possibilities: Prospects for Internet Democracy. Conference Papers, "Internet, Democracy, and Public Goods," Belo Horizonte, Brazil, November 6-10, 2000.

Fruchterman, T. and Reingold, E (1991). "Graph Drawing by Force-Directed replacement." *Software--Practice and Experience* 21:1129-1164.

Gastil, J. (2000). Is Face-to-Face Citizen Deliberation a Luxury or a Necessity? *Political Communication*, 17, 357-361

Gutmann, A. and Thompson, D. (1996). *Democracy and Disagreement.* Cambridge, MA: Harvard University Press.

Habermas, J. (1984). *The Theory of Communicative Action, Vol. 1.* (T. McCarthy, Trans.). Boston, MA: Beacon.

(1991) The Public Sphere. In C. Mukerji & M. Schudson (Eds.), *Rethinking popular culture*. Berkeley: University of California Press.

Hardin, R. (1999). Deliberation: Method, not Theory. In Macedo, S. (ed.), *Deliberative Politics: Essays on Democracy and Disagreement*, 70-87.

Klein, H. (1999). Tocqueville in Cyberspace: Using the Internet for Citizen Associations. *The Information Society*, 15, 213-220.

McBurney, P. and Parsons, S. (2001). Intelligent systems to support deliberative democracy in environmental regulation. *Information and Communications Technology Law,* 10 (1): 33-43.

Noelle-Neumann, E. (1984). *The spiral of silence: Public opinion – our social skin.* Chicago: University of Chicago Press.

Papacharissi, Z. (2002). The virtual sphere: The Internet as a public sphere. *New Media and Society*, 4, 9-27

Payne, R. (1996). Deliberating Global Environmental Politics. *Journal of Peace Research*, 33 (2), 129-136.

Price, V. and Capella, J. (2001). Online Deliberation and its Influence: The Electronic Dialogue Project in Campaign 2000. Conference Papers, annual meeting of the American Association for Public Opinion Research, May 2001.

Resnick, P. (2000) Beyond Bowling Together: SocioTechnical Capital. In J. Carroll, *HCI in the New Millennium.* New York: Addison Wesley, Chapter 29.

Sanders, L. (1997). Against Deliberation. *Political Theory.* 25(3), 347-364.

Shapiro, I. (1999). Enough of Deliberation: Politics is about Interests and Power. In Macedo, S. (ed.), *Deliberative Politics: Essays on Democracy and Disagreement*, 28-38.

Warren, M. (2002). Deliberative Democracy. In Carter, A. and Stokes, G. (eds.), *Democratic Theory Today.* Malden, MA: Polity Press. 173-202

Witschge, T. (2002). Online Deliberation: Possibilities of the Internet for Deliberative Democracy. Conference papers, "Prospects for Electronic Democracy," Carnegie Mellon University.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion.* New York: Cambridge University Press.

CT+M

An Amplification of Sensationalism:

Comparing the Tonal Values of the *New York Times* to the *New York Post* Using Whissell's Dictionary of Affect in Language

Paul Crandon, Ph.D.
University of North Alabama

John J. Lombardi, Ph.D.
Frostburg State University

> *"Choices of words and their organization into news stories are not trivial matters. They hold great power in setting the context for debate, defining issues under consideration, summoning a variety of mental representations, and providing the basic tools to discuss the issues at hand."*
>
> **– Pan & Kosicki**[1]

Introduction

This brief investigation marks an attempt to borrow what we believe is a clever and useful research instrument/method and apply it to a practical discussion in journalism and media studies. The instrument is a database of thousands of words scale-rated in emotional dimensions, and the discussion at hand is how to define and measure sensationalism.

The idea centers around the connotative meanings – or affective tonal value – of words, and the idea that this tone can be measured and dealt with statistically to compare and content analyze bodies of text.

While many words share roughly the same *denotative* meaning, e.g. wallet and billfold, no two words convey the exact same affective, or *connotative*, meaning; all words evoke emotional responses that are different from all other words. Though there is some variance between respondents – not all people react to a word in the exact same manner – scales have been developed that offer measures of emotional tone, and they have received at least some evidence of external validity. What if there were a corpus of words, thousands of them, that were each rated in multiple dimensions of emotion, along with computer programming that can render emotional measures from any text, large or small, instantly? There is.

---

[1] **Z. Pan and G. M. Kosicki,** "Framing analysis: An approach to news discourse," *Political Communication*, 10 (1993): 70.

One such instrument was developed by Cynthia Whissell, called the *Dictionary of Affect in Language*, which has been used in conjunction with computerized content analysis software to measure the affective tone of copy from a host of sources. Although Whissell's is not the first attempt to catalogue the affective element of large numbers of words, the DAL is the most comprehensive and extensively used to date.[2]

The dictionary was composed using Osgood's semantic differential techniques[3] to rate thousands of words in terms of three important dimensions: the words' pleasantness (pleasant – unpleasant), activation (active – passive), and the words' imagery (hard – easy to imagine). The goal was to compile a reference list of the affective or emotional "meanings" of frequently used words that could later be used to analyze text by computer. Words were chosen for inclusion based on their frequency of use in common spoken and written English. In the end, nearly 10,000 words were checked for spelling and included in the list.[4]

The usefulness of such an instrument should be quite apparent: researchers could use Whissell's dictionary to measure the tone of large quantities of copy instantly, comparing publications alone and to each other, and across time. Studies could use these methods to examine the emotional tone with which a particular issue is portrayed by different media and whether that tone changes over time. Political speeches could be analyzed for changes in emotion from one to the next or from speaker to speaker. One could compare the tone of coverage from local media versus national media, for example, or analyze coverage from a single source over the life of an issue. Studies in public relations could look at the tonal values of an in-house newsletter compared with mainstream media (Are newsletters more pleasant than "real" news? Less active? Higher in imagery?). Advertising scholars and executives alike could examine trends in the field and study the efficacy of ads using different tonal values. Provocative questions could be probed – How has coverage of AIDS changed in tone from the early 80s to today? Are the news media becoming more arousing in their coverage? Less arousing? Does coverage of the War in Iraq differ in tone from coverage of the Gulf War, the Vietnam War or other military actions?

---

[2] **K. Sweeney and C. Whissell**, "A dictionary of affect in language: I. Establishment and preliminary validation," *Perceptual and Motor Skills*, 59 (1984): 695-698. **C. Whissell**, "Pleasure and activation revisited: Dimensions underlying semantic responses to fifty randomly selected 'emotional' words," *Perceptual and Motor Skills*, 53 (1981): 871-874. **C. Whissell and K. Charuk**, "A dictionary of affect in language: II. Word inclusion and additional validation," *Perceptual and Motor Skills*, 61 (1985): 65-66. David Heise also constructed earlier profiles of 1,000 common English words in **D.R. Heise**, "Semantic differential profiles for 1000 most frequent English words," *Psychological Monographs: General and Applied*, 79:8 (1965): 1-31.

[3] **C. E. Osgood, G. J. Suci, and P. H. Tannenbaum**, *Measurement of Meaning* (Urbana: University of Illinois Press, 1957).

[4] **K. Sweeney and C. M. Whissell**, (1984): 696.

Using the DAL

Each word in the dictionary has a decimal number rating between 1-3 for each of the three scales of PLEASANTNESS, ACTIVITY, and IMAGERY. A body of text can be computer analyzed and a mean rating for each dimension can be found. For example, the word "yesterday" was rated 2.57 on the pleasantness scale, 1.83 on the activity scale, and 1.60 on the imagery scale. This would indicate that subjects found this word to be relatively pleasant, not particularly active (or passive), and somewhat difficult to imagine. In another example, this time a much more neutral word, the word "central" scored as follows: 1.67 pleasantness, 1.67 activity, and 1.40 imagery. It is easy to see that subjects found this word to be neither pleasant nor unpleasant, neither active nor passive, and perhaps a bit difficult to imagine. With **three separate scores** for thousands of commonly used words, one can begin to appreciate the utility in the dictionary.

In addition, Whissell has devised a method of scrutiny whereby extreme words can be located and tabulated. Words in the extremes of these three dimensions have been isolated and given appropriate labels. For example, decidedly PLEASANT words include those words that rated in the 10 percentile of pleasantness of all rated words. Similarly, UNPLEASANT words are those words that were rated in the bottom 10 percentile of this dimension. ACTIVE words are those words rated by subjects in the top 10 percentile of the activity dimension, and PASSIVE words are words in the bottom 10 percentile of this dimension. Finally, HIGH IMAGERY and LOW IMAGERY words are those words that scored in the top and bottom 10 percentiles of this dimension. Thus far, six different categories of extreme words have been "tagged" in the dictionary. Notice that in each grouping, the line of demarcation was located at 10 percent.

Whissell has also combine two of these dimensions – pleasantness and activity – to form four more categories of extreme words. By taking the top and bottom quartiles of each of these, Whissell devised these new categories: NICE words (top 25 percentile for pleasantness/bottom 25 percentile for activity), SAD words (bottom 25 percentile for both pleasantness and activity), CHEERFUL words (top 25 percentile of both pleasantness and activity), and NASTY words (bottom 25 percentile for pleasantness and top 25 percent for activity). Notice that when two dimensions are combined, the range of inclusion is broadened to 25 percent for each scale.

Whissell's dictionary has been used in a number of unique studies, most merely designed to test the fitness of the instrument itself. In a stylometric study examining the song lyrics of Paul McCartney and John Lennon, for example, Whissell was able to replicate earlier critical studies whose findings suggest which writer was more happy, cheerful, etc. and which was more sad or depressed. She was also able to show quantitatively how the "mood" of the authors' lyrics changed over time, again, in agreement with other literary and music scholars' previous qualitative or hand-coded works. In addition, the dictionary was able to take a sample of song lyrics and correctly identify which writer composed it based primarily on the tone of the passage.[5] This is an important finding because it lends credibility to Whissell's methods and instrumentation, and offers the DAL as a valid tool for stylometrists. Others have used the instrument to explore a number of issues across a host of disciplines, from measuring the emotional tone of open-ended responses in management questionnaires[6] to comparing the written sexual fantasies of men and women.[7]

Whissell herself has ventured into the realm of media studies. In one example, Whissell and McCall found differences in the tonal values of advertisements aimed at men and women. The authors compared the copy from print ads in leading men's magazines such as *Gentlemen's Quarterly* and *Popular Mechanics* to those found in women's magazines, such as *Ladies' Home Journal* and *Homemakers.* The study found that ads directed at men were more arousing and less pleasant than the ads aimed at women.[8] Within this study, a follow-up experiment revealed that women tend to rate ads as more successful in their appeals when words higher in pleasantness are used, while ads using words higher in arousal were rated more effective by both men and women. This study was later extended to incorporate the third dimension – imagery – and to include children as subjects as well.[9]

---

[5] **C. Whissell,** (1996).

[6] **K. W. Mossholder, R. P. Settoon, S. G. Harris, and A. Armenakis,** "Measuring emotion in open-ended survey responses: an application of textual data analysis," Journal of Management, 21:2 (Summer, 1995) 335-355.

[7] **Stephanie L. Dubois,** "Gender differences in the emotional tone of written sexual fantasies." *The Canadian Journal of Sexuality*, 6:4 (Winter, 1997) 307-315.

[8] **C. Whissell and L. McCall** (1997): 365

What seems noteworthy for this discussion and for communication scholars, is the fact that these studies have not found their way into our journals. Read on:

In an experiment designed to investigate the emotional tone of newspaper headlines, Fournier *et al.* sought to provide a testable operational definition of "sensationalism" using the Whissell DAL.[10] The researchers obtained newspaper headline copy from three newspapers: the *Toronto Globe and Mail* and the *Wall Street Journal*, and one considered to be sensational, the *Toronto Star*. As an external check, the authors also included a similar sampling of article titles from the academic journal *Psychological Reports*. Results indicate that by using the DAL to analyze copy the researchers were able to identify copy deemed sensational:

> Sensationalism could be defined, in terms of the Dictionary of Affect, in one of two ways: it could involve a high level of activity [arousal levels] in language regardless of evaluation [pleasantness] in which case *Toronto Star* headlines and *Psychological Reports* titles would both be classified as "sensational". Although readers of the *Toronto Star* might readily agree with this classification, authors of papers in *Psychological Reports* would probably be surprised to find their material so described. An alternative definition of sensationalism would require the relatively high usage of active, unpleasant words. By this definition, titles in the *Toronto Star* would still be classified as sensational, but those in *Psychological Reports* would not.[11]

Some irony (at least for media scholars) in this case might be found in the fact that this was published in a psychology journal.

These examples are among a limited number of studies attempting to shed other than anecdotal light on sensationalism. Few attempts have been made to operationally define sensationalism, let alone quantify or measure it. Indeed, an informal database search of 20 years of refereed journals, using "sensationalism" as the sole subject search term with no other limitations, yielded only 82 articles, with most of these (53) in unrelated fields.[12] Of those appearing within the fields of communication/media (29 total), 18 appeared in the *American Journalism Review*. Only one article attempted to clarify our conception of sensationalism beyond the popular usage of the term: Grabe, et al., did study the packaging of TV news and how video maneuvers (zooms in/out, shot duration, etc.) and decorative effects (sound effects, wipes and fades, etc.) associated strongly with what has come to be called sensational TV journalism.[13]

---

[9] **L. Rovinelli and C. Whissell,** " Emotion and style in 30-second television advertisements targeted at men, women, boys, and girls," *Perceptual and Motor Skills*, 86:3 (1998) 1048-50.

[10] **M. Fournier, M. Dewson, and C. Whissell,** "The Dictionary of Affect in Language: VI. 'Sensationalism' defined in terms of affective tone," *Perceptual and Motor Skills*, 63 (1986): 1073-1074.

[11] Ibid., 1074. This study was conducted without the benefit of the third dimension – IMAGERY – which was later

Research Questions

How does copy taken from an arguably non-sensational newspaper differ from copy taken from a newspaper that is arguably sensational? Can we arrive at a better understanding of what constitutes sensationalism by examining the features of copy from a source that has been called sensational? Can some validation of the instrument itself be extended, should results intuitively make sense?

Method

To address these research questions we analyzed the headlines and the lead sentences from three weeks of front-page lead stories from both the *New York Times* and the *New York Post*. These publications were selected because of their historically competing approaches to news reporting. The *New York Times* is often referred to as the nation's "newspaper of record" and enjoys a long-standing reputation for high quality reporting. The *New York Post*, by comparison, tends to be perceived as the "yin" to the *Times*' "yang," with less concern over journalistic integrity.[14] In an effort to avoid the potential for repetition of weekday stories often found in weekend editions, only front-page stories from the Monday through Friday

---

incorporated into the Dictionary.

[12] **InfoTrac** searched April 1, 2003. These groupings (followed by number of "hits") emerged: Science/nature (7), criminal/law (7), medicine (4), sociology/political science (10), anthropology/cultural studies (6), theatre/film (5), literature (5), business (3), history (2), and other/unspecified (4).

[13] **M. E. Grabe, S. Zhou, and B. Barnett**, "Explicating sensationalism in television news: content and the bells and whistles of form," *Journal of Broadcasting & Electronic Media*, 45:4 (Fall, 2001) 635-55.

[14] Prominent journalism historian **Mitchell Stephens** wrote an essay on sensationalism, specifically featuring Rupert Murdoch and the *New York Post* ("Viewpoints: About Rupert Murdoch," *Newsday*, April 2, 1993, p. 52); Harry Stein specifically contrasted the *Times* and the *Post*, calling the latter "subtle as a Vegas floor show" when it comes to tabloid journalism ("New York's Tabloid Treasure," *Urbanities*, 11:1 (2001).

editions of each publication were included in this study. Page-one lead stories (one from each day) from February 24 through March 21, 2003 were obtained through a LexisNexis database search, yielding **four sample data sets**: 15 headlines from the *Times* and 15 from the *Post*, and as many leads from each of the two papers.

The *New York Post* routinely has only one major story on its front page, attributable largely to its tabloid format. This made locating this as the lead story effortless. The *New York Times*, however, obviously prints several stories on the front page of each edition, so some judgement had to be made to determine which would constitute the lead story. The lead was determined to be the story located in the first column, above the fold, in the upper left portion of each edition. Once the lead stories were identified by a visual inspection of the paper version, the digital text was downloaded using LexisNexis. Using this database retrieval method eliminates the risk of using "early" or regional editions (only final national versions appear in the database), as well as the risks of contamination associated with OCR software.

Once the text for each story's headline and lead was obtained, separate sample files were created and each file was checked for any spelling errors or extraneous wording that may have been copied over from the Nexis download. These files were loaded directly into software designed exclusively for use with Whissell's dictionary for analysis.[15]

Results

All told, 15 headlines and 15 leads were examined from each of the two publications. With respect to headlines, the *New York Times* tallied 164 words or 10.93 words per headline while the *New York Post* tallied 188 words or 12.53 words per headline (SEE Table 1). Keeping in mind that the DAL does not recognize all words in the English language it is important to realize that 98.0% of all words in the *Times*' headlines and 87.5% of all words in the *Post's* headlines were recognized by the DAL.

A total of 566 words or 37.73 words per sentence were found in the *New York Times* leads. A total

---

[15] The software, simply called *Whissell's Dictionary of Affect in Language*, was developed by Dr. Paul Duhamel, a former graduate student under Dr. Whissell. Copyright 1998-2002 Human Development Consulting – info@hdcus.com.

TABLE 1 – FREQUENCIES

	NY Times	NY Post
Headlines		
Total words	164	188
Average length (words)	10.93	12.53
Hit Rate	98.0%	87.5%
Leads		
Total words	566	623
Average length (words)	37.73	41.53
Hit Rate	96.6%	94.4%

of 623 words or 41.53 words per sentence were found in the *New York Post* leads. A total of 96.6% (*Times*) and 94.4% (*Post*) of all words in the leads were analyzed by the DAL (SEE Table 1).

With respect to headlines, significant differences were found in PLEASANTNESS between the two newspapers (f = 6.30, df = 195, p = .013). While the means of the two samples suggest little differences in the arithmetic average (1.74 versus 1.81), the higher standard deviation seen in *Post* headlines implies a greater use of extreme words (regardless of whether they fall within the 10 percentile necessary for classification as a decidedly pleasant word). No statistical significance was found when comparing the mean level of ACTIVATION or IMAGERY among headlines. However, chi square analysis did reveal some significant differences among the percentage of certain classes of extreme words used in the headlines. For example, the *Post* used significantly more PLEASANT words than did the *Times* (chi square = 9.67, df = 1, p = .002) (SEE Table 2).

Chi square analysis also revealed that the *Post* used significantly more HIGH IMAGERY words in its headlines than did the *Times*. These words are ones that fall into the top 10% of the IMAGERY category (chi square = 3.81, df = 1, p = .051). Of additional significance, it was found that the *Times* used significantly more SAD words than did the *Post* (chi square = 4.75, df = 1, p = .029). According to the DAL, SAD words are ones that score at the bottom 10% of both ACTIVATION and PLEASANTNESS (SEE Table 2).

With respect to the crafting of lead sentences, another t-test reveals significant differences in their use of PLEASANT words (f = 13.01, df = 766, p = .000). The *Post* also had a significantly higher mean IMAGERY rating than did the *Times* (f = 8.22, df = 766, p = .004) (SEE Table 3).

Table 3 further shows that when we break down the words used in the headlines into each of the ten extreme word classes, there are several significant differences between the two publications. The *Post*

TABLE 2 – HEADLINES

	NY Times (σ)	NY Post (σ)			
Interval Data (t-test)			**f**	**df**	**Sig.**
Mean Pleasantness	1.74 (.39)	1.81 (.51)	6.30	195	.013
Mean Activation (Arousal)	1.83 (.41)	1.86 (.45)	.803	195	.371
Mean Imagery	1.99 (.64)	2.10 (.71)	2.28	195	.133
Nominal data (Chi-Square)			**X^2**	**df**	**Sig.**
% Pleasant	2.04	14.14	9.67	1	.002
% Unpleasant	15.31	14.14	.053	1	.818
% Active	7.14	11.11	.934	1	.334
% Passive	15.31	13.13	.191	1	.662
% High Imagery	10.20	20.20	3.81	1	.051
% Low Imagery	13.27	13.13	.001	1	.978
% Nasty (hybrid)	10.20	12.12	.182	1	.669
% Sad (hybrid)	7.14	1.01	4.75	1	.029
% Nice (hybrid)	1.02	4.04	1.82	1	.178
% Cheerful (hybrid)	7.14	9.09	.250	1	.617

TABLE 3 – LEADS

	NY Times (σ)	NY Post (σ)			
Interval Data (t-test)			**f**	**df**	**Sig.**
Mean Pleasantness	1.81 (.34)	1.82 (.41)	13.01	766	.000
Mean Activation	1.71 (.42)	1.72 (.40)	1.33	766	.249
Mean Imagery	1.70 (.67)	1.84 (.74)	8.22	766	.004
Nominal data (Chi-Square)			**X^2**	**df**	**Sig.**
% Pleasant	4.63	8.18	4.06	1	.044
% Unpleasant	4.88	8.71	4.44	1	.035
% Active	6.43	6.33	.003	1	.957
% Passive	24.16	20.32	1.64	1	.200
% High Imagery	7.46	12.66	5.78	1	.016
% Low Imagery	30.85	26.65	1.65	1	.199
% Nasty	4.88	8.97	4.99	1	.025
% Sad	4.37	4.49	.006	1	.938
% Nice	1.03	4.22	7.72	1	.005
% Cheerful	5.66	5.01	.157	1	.692

used significantly more PLEASANT words in its leads than did the *Times* (chi square = 4.06, df = 1, p = .044). The *Post* used significantly more UNPLEASANT words, as well (chi square = 4.44, df = 1, p = .035). The same can be said for HIGH IMAGERY words, where 12.66% of words used in the *Post* leads were considered HIGH IMAGERY words, compared to only 7.46% of words in the *Times* leads (chi square = 5.78, df = 1, p = .016). *Post* leads also contained a higher percentage of NASTY words and NICE words than did the *Times* (8.97% vs. 4.88% for nasty and 4.22% vs. 1.03% for nice). Both categories proved to be significant. NASTY words were significantly higher in the *Post* leads (chi square = 4.99, df = 1, p = .025) as were NICE words (chi square = 7.72, df = 1, p = .005) (SEE Table 3).

Discussion

Before delving too deeply into this discussion we must first preface our remarks with a quick definition of "sensationalism." The concept of sensationalism as it pertains to journalistic practices carries a decidedly negative connotation. What should be considered when thinking of sensationalism as a descriptive term is that the essence of the word connotes an altering of the senses, regardless of evaluation. In this case the "sense" is not touch, taste, sight, sound, or smell by itself, but is a combination of any or all of these senses. The sense that is jogged by sensationalism is that of the imagination.

Perhaps one of the most provocative findings presented here lay in the counterintuitive significant differences found in the overall pleasantness of the "sensational" copy. Indeed, the headlines from the source assumed to be sensational were found to be significantly more PLEASANT in overall tone, and were also found to incorporate significantly more decidedly PLEASANT words. There were also significantly more NICE words included in the *Post* leads than in the *Times*. At first blush, this may seem to contradict the conventional characterization of sensationalism: that sensational journalism feeds only on the negative, that sensational journalists represent the bottom-dwellers, reporting only the seedy underbelly of bad news. On the other hand, could not what we might call sensational in other contexts be significantly more pleasant? Would we not say something is "sensational" if it appeals to senses other than the baser ones, such as a "sensational" Broadway musical? As an exploratory study, there was the more or less tacit agreement that if there were differences to be found in pleasantness, they would in all likelihood be in the other direction. That this isn't the case tells us perhaps as much about our preconceived ideas of sensationalism as it does about what sensationalism might really involve.

This is not to imply a full vindication of the *Post*, however. The *Post* leads boasted significantly more SAD words in their headlines, as well as significantly more UNPLEASANT and NASTY words in their leads. It seems that indeed sensationalism can operate in both directions. It appears the *Times* may have adopted more of a keep-it-in-the-middle approach to their reporting, while the *Post* seems to feel free to steer to both sides of the road. These findings seem to support a re-tooling of some of the accepted wisdoms surrounding sensational journalism. It seems that sensationalism may merely be the presence of more extremes, be they NASTY and UNPLEASANT, or NICE and PLEASANT.

Another finding that deserves attention is that of IMAGERY. There were nearly twice as many HIGH IMAGERY words used in the *Post* headlines, and 70 percent more in the leads. Indeed, the overall tone of the leads under this dimension was significantly higher for the *Post*, accompanied with greater variance (the tendency to lean toward the extremes) as well. This falls comfortably under the popular ideas of what constitutes sensational reporting. It appears from these data that the increased use of words that are more concrete, or easier to imagine, words helps produce what might "feel" like sensationalism. It should be reiterated, however, that this spike in IMAGERY still does not suggest a direction of pleasantness, nor does it carry any necessarily negative connotations. This simply suggests that using higher IMAGERY words – positive or negative – contribute to what is popularly known as sensationalism. Does this imply that editors should advise cub reporters to stay away from high IMAGERY words, lest they be accused of sensationalizing? Perhaps that question is best left as rhetorical. But consider this: How might the stodgy stalwart journalists of just a generation or two ago respond to the brassy new journalism of today? Would they consider the leads we see in contemporary news coverage to be sensational? Again, we'll choose to leave this question as rhetorical.

The principal goal of this paper was to attempt to apply Whissell's dictionary as a methodology to media studies, particularly as it might elucidate research on sensationalism. To this end, we wondered whether two arguably disparate publications would manifest differences in affective tonal value. If a study could demonstrate this trend, it would offer some support for the methodological value of Whissell's Dictionary of Affect in Language.

Being able to measure and compare affective tonal values marks a keen methodological advancement, and this research gives some credence to the idea that an affective tone can be identified and measured within a text, and compared with valid results to that of other compositional bodies. Part of the reason for undertaking this project was to test the efficacy of using Whissell's DAL to measure the affective elements of a news story. Whissell's dictionary represents years of research, numerous hours of test administration and coding, and a host of studies in search of cross-validation and internal methodological rigor. These efforts have produced a list nearly 10,000 words long, each with a score for three dimensions – PLEASANTNESS, ACTIVATION, and IMAGERY – and represent the continuation of the work started by Osgood and others a half century ago. There is corroborating evidence that seems to validate the instrument; it was repeatedly able to replicate by computer results that were previous found critically or by hand. Stylistic matters relating to tone and to word and sentence length can distinguish one author from another. Differences in advertising copy were found using the DAL that made good sense intuitively (that advertising aimed at men tends to be more arousing and less pleasant).

However, one drawback to Whissell's work is perhaps the lack of attention it has received outside its own niche; few if any scholars outside Whissell's group have tested the efficacy of the DAL. Certainly it is beneficial to have scholars from other areas and backgrounds such as communication and sociology investigate these new measures and techniques independently. This study is one attempt to begin the process of assimilation of this work from one discipline into another.

The extreme words that are outlined in the DAL are categorized as PLEASANT, UNPLEASANT, ACTIVE, PASSIVE, HIGH IMAGERY, LOW IMAGERY, NASTY, SAD, NICE, and CHEERFUL. This research has found that the *New York Post* more often than the *New York Times* uses higher amounts of these words. A quick review of the results suggests that while the *Times* used significantly more sad words in their headlines than did the *Post*, the *Post* used significantly more pleasant and high imagery words in their headlines. The *Post* also used significantly more pleasant, unpleasant, high imagery, nasty, and nice words in their leads while the *Times* did not use any of these types of words significantly more than the *Post*.

CT&M

Running Head: UNIDIMENSIONALITY

Do Mass Communication Studies Test Measures for Unidimensionality?

John D. Richardson and Frederick Fico

Michigan State University

Address Correspondence to:

John D. Richardson
Michigan State University
Department of Telecommunication
419 Communication Arts & Science Bldg.
East Lansing, MI 48824-1212
Tel: 517-432-1334
Fax: 517-355-1292
richa377@msu.edu

Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication, Kansas City, MO (July, 2003). John D. Richardson is a doctoral candidate in Mass Media Studies at Michigan State University, where Frederick Fico is a Professor in the School of Journalism.

Abstract

A review of leading mass communication journals indicates that studies introducing measures of latent constructs rarely considered unidimensionality, a critical element of construct validity. Drawing on psychometric literature, it is demonstrated that procedures commonly used to assess/develop mass communication measures, particularly Cronbach's alpha and exploratory factor analysis, do not examine whether a measure is unidimensional. Moreover, the review of prior studies also suggests widespread use of Kaiser's "Little Jiffy," a contraindicated combination of exploratory factor analytic techniques. Psychometric research explicating problems with these techniques is reviewed. Use of confirmatory factor analysis to test unidimensionality is also discussed.

Do Mass Communication Studies Test Measures for Unidimensionality?

Introduction

Mass communication researchers often seek to assess abstract phenomena (e.g., newspaper credibility, viewer involvement) that are not directly observable. Single-item indicators are an inferior approach: they are highly prone to measurement error and the construct of interest is rarely synonymous with any single indicator. Therefore, common practice is to administer a cluster of items, sum or average the responses, and treat the composite score as a measure of the underlying construct.

However, if such a measure lacks unidimensionality, a score derived from responses to the cluster of items is not interpretable. A measure is *unidimensional* where an underlying measurement model specifies that one and only one underlying construct (and error) is causally antecedent to each item in the cluster (Hunter & Gerbing, 1982).

Our review of the mass communication literature suggests that studies that develop and/or assess latent measures usually do not test their unidimensionality. In addition, our review also suggests that exploratory factor analyses often employ a combination of techniques, known as Kaiser's "Little Jiffy" (Kaiser, 1960, 1970), contraindicated by psychometric research. The goal of this paper is to stimulate mass communication researchers to rethink their techniques for building measures of mass communication constructs.

Unidimensional Constructs

A *construct* is a conceptual term invented for the purpose of describing a phenomenon of theoretical interest (Edwards & Baggozzi, 2000). Constructs themselves are not substantively real, but they refer to phenomena believed to exist in reality. The creation of a construct gives rise to several issues, including: (a) whether the tentative list of indicators needs modification and/or extension; and (b) the relationship between each indicator and the construct (Hogan & Nicholson, 1988).

Theoretical models are concerned with relationships between phenomena of interest. However, empirical studies do not assess constructs directly. Instead, an empirical measure serves to indirectly estimate a phenomenon assumed to be associated with the construct of interest. A *measure* is an observed score gathered via self-report, interview, observation, or other empirical means. A measure is intended to serve as an analog to a construct. *Indicators* refer to scores on individual items that, when cumulated to form an index or scale, comprise a measure (Edwards & Baggozzi, 2000).

Covariation among empirical measures facilitates theoretical interpretation only if the measures are valid tests of the phenomena presumed to be associated with the constructs (Hunter & Gerbing, 1982). Ozer (1999) suggested that construct validity is most accurately conceived of not as a quality of the measure, but as a characteristic of the relationship inferred between the measure and the phenomenon purportedly assessed. The process of validating a measure is dialectical; the researcher must constantly move back and forth between two issues: (1) Does the theoretical phenomenon the construct attempts to describe exist in fact? If so: (2) Do the empirically observed responses assess changes or differences in that phenomenon? (Hogan & Nicholson, 1988).

The relationship between any observed measure (including valid measures) and the phenomenon of theoretical interest is imperfect due to measurement error. To mitigate this problem, multiple indicators are employed. If the indicators are all systematically influenced by only one construct (i.e., if the measure is unidimensional), the influence of the construct on responses to indicators is common to them all, but the processes creating errors will vary randomly from one indicator to the next. Accordingly, the common thread running through multiple indicators should be more indicative of the phenomenon underlying the construct. Hence, unidimensionality is a crucial element of construct validity. Unidimensionality may be regarded as a falsifiable hypothesis requiring explication and testing. A measurement model explicates this hypothesis (Hunter & Gerbing, 1982).

Measurement Models

The classical measurement model conceptualizes a construct as an independent variable and its indicators as dependent variables. Figure 1, adapted from West's (1994) study of Meyer's (1988) newspaper credibility scale, provides an example of this classical model. Meyer conceptualized newspaper credibility as the extent to which readers believe a newspaper. West's measurement model hypothesized that the construct, newspaper credibility, is the only independent variable in the system, and that five indicators (Fair, Unbiased, Tells the Whole Story, Accurate, and Can be Trusted) are all dependent on the newspaper credibility construct.

INSERT FIGURE 1 ABOUT HERE

Several corollaries flow from these hypotheses. First, because newspaper credibility is the only independent variable, the model implies that the measure is unidimensional. Second, the model also implies that each indicator is positively correlated with every other indicator. Third, the model implies that inter-indicator correlations result solely from the fact that each indicator shares common variance with the construct. In addition, an incomplete sample of indicators can serve as a valid measure of the construct. In other words, if one of Meyer's (1988) five indicators were eliminated, the resultant four-indicator measure would be just as valid (although it might be less reliable), because each indicator assesses newspaper credibility and nothing else.

Bollen and Lennox (1991) described another measurement model in which causation flows in the opposite direction, i.e., from the indicators to the construct. Edwards and Bagozzi (2000) used the term *formative model* to describe this latter type. As an example, Bollen (2002) conceptualized exposure to media violence as a dependent construct emerging from the cumulative influence of three indicators: time spent attending to violent content in television, in movies, and in video games. By adding time spent attending to violent content in music and in print, we arrive at a graphical depiction of a violent media exposure model, shown in Figure 2.

INSERT FIGURE 2 ABOUT HERE

Unlike the classical model, the formative model posits no hypotheses as to the cause(s) of the indicators. The indicators are not driven by a unidimensional construct, but instead are separate dimensions that define the construct. Correlations among the indicators may be positive, negative, or zero.[1] Failure to include any unique, significantly correlated indicator omits an element of the construct, thereby impairing the validity of the measure.

Determination of the causal order of indicators and constructs, i.e., whether a "classic" or "formative" model is most appropriate, might in some cases present a "chicken and egg" dilemma. In general, where a change in the construct is expected to result in a simultaneous change in all indicators, the classical measurement model is implied. Conversely, where a change in an indicator would be expected to result in a change in the construct, even while other indicators remain constant, then this is more suggestive of a formative measurement model (Bollen & Ting, 2000).

The majority of social science measures treat indicators as dependent upon constructs, in accordance with the classical measurement model (Bollen & Ting, 2000). Most measures of mass communication constructs are also probably conceptualized in this manner (albeit by implication only). However, our review of studies in which measures of mass communication constructs were introduced found that unidimensionality was rarely tested.

Review of Mass Communication Studies

A review of mass communication studies was performed to assess the procedures most commonly used in the development of measures of mass communication constructs. Studies reviewed were published in *Mass Communication & Society* from 1998 (the date of its inception) to Spring 2002, *Journalism & Mass Communication Quarterly* from 1992 to Spring 2002, and *Newspaper Research Journal* from 1992 to Spring /Summer 2002. We found a total of 124 studies that introduced one or more new measures of a mass communication construct. The main findings are set forth in Table 1.

INSERT TABLES 1 AND 2 ABOUT HERE

Results indicated that most mass communication studies retain a new measure so long as Cronbach's alpha surpasses some benchmark, typically .70. Indeed, this is somewhat of an overstatement; several published studies employed measures with lower alpha values. A large minority of studies used exploratory factor analytic techniques to derive measures *ad hoc*. Confirmatory factor analysis to test an a priori measurement model was rarely performed.

As Table 2 demonstrates, studies that employ exploratory factor analysis frequently failed to report which techniques were employed. Among studies that did specify the exploratory factor analytic techniques used, principal components analysis was the most common factor extraction method and varimax the most common factor rotation method. Eigenvalues-greater-than-one was frequently used to decide the number of factors to retain.

These findings suggest many mass communication studies use Cronbach's alpha and/or exploratory factor analysis to establish the "usability" of a measure. In the sections that follow, we review psychometric evidence to illuminate some limitations of these techniques.

Cronbach's Alpha

The procedure most commonly followed by mass communication researchers when introducing a new measure seems to be as follows: (1) assemble items that appear to possess face validity; (2) determine if Cronbach's alpha exceeds some "standard of acceptability," and if so; (3) treat the measure as though it were "usable." However, this procedure ignores the all-important issue of unidimensionality.

First, consider that the classical measurement model posits unidimensionality, i.e., there must be one and only one latent construct antecedent to each indicator. Cronbach's alpha *presumes* that the measure is unidimensional, it does not assess unidimensionality (Gorsuch, 1997; Green et al., 1977; Hattie, 1985; McDonald, 1981). Indeed, Green and colleagues demonstrated that a set of indicators influenced by five distinct constructs could yield an alpha of .81, (an alpha value generally regarded as very "acceptable"). In other words, a measure that is not unidimensional can nevertheless yield a high alpha.

Next, consider a construct premised upon the formative measurement model. If some indicators are less reliable than others, the value of Cronbach's alpha decreases. This could possibly induce the researcher to drop less reliable indicators to obtain a higher alpha. However, in order for a formative measure to validly assess the construct of interest, all unique indicators significantly correlated with the construct must be included, regardless of their reliability relative to other indicators. Dropping indicator(s) to obtain a higher alpha can result in the omission of an important aspect of the construct (Bollen & Lennox, 1991).

Exploratory Factor Analysis

The next most common method employed by mass communication researchers who develop a new measure is some form of exploratory factor analysis (EFA). The goal of EFA is to move toward a more parsimonious conceptual understanding by attempting to identify the fewest abstract constructs that explain the common variance among the measured indicators. However, EFA is a data-driven, *post hoc* approach. This is in stark contrast to confirmatory factor analysis (discussed below), in which a hypothesis is formed before data are gathered, and the data are analyzed to test the hypothesis. Accordingly, EFA results should never be interpreted as evidence of the unidimensionality of a measure (Borgatta et al., 1986; Fabrigar et al., 1999; Floyd & Widaman, 1995; Gorsuch, 1997; Loehlin, 1998; Reise et al., 2000). Theories positing relationships between constructs "identified" by EFA alone should be regarded as highly preliminary.

Furthermore, our review of mass communication studies demonstrated widespread use of a three-step approach to EFA: (1) extract factors via principal components analysis (PCA); (2) retain factors that have an eigenvalue greater than one (E1); and (3) rotate factors using varimax rotation. Known as Kaiser's "Little Jiffy," this combination of techniques is contraindicated.[2]

Principal Components Analysis

PCA is a factor extraction method based on the formative model of measurement, not the classical measurement model. The resultant principal components are emergent variables, dependent on the indicators (Loehlin, 1998). Accordingly, it is theoretically illogical to use the PCA method of factor extraction when attempting to develop a measure that will follow (even if only implicitly) the classical measurement model. Several psychometricians have suggested the default settings on popular software packages influence many unwitting researchers to use PCA (Fabrigar, et al., 1999; Gorsuch, 1990; Reise et al., 2000). For example, PCA is the default factor extraction method on Statistical Package for the Social Sciences (SPSS). Hence, perhaps it is little wonder so many mass communication studies fail to report which method of factor extraction was selected (or when they do report it, offer no rationale for selecting PCA)—very often the researcher did not select an extraction method—they let the software choose it for them! Nevertheless, a minority of psychometricians assert that PCA is an acceptable option, arguing that PCA yields results similar to other factor extraction methods so long as the measure includes a "reasonable" number of indicators (Velicer & Jackson, 1990).

Principal axis factoring (PAF) is the EFA factor extraction method most commonly used in psychology research, and is based on the classical measurement model. Snook and Gorsuch (1989) found that compared to PAF, PCA resulted in inflated factor loadings, especially where there were few variables and low factor loadings. Even with 36 measured variables and factor loadings of 0.80, the differences between PCA and PAF remained significant. Overall, PAF yielded results that more accurately reflected the known population parameters. PCA tended to present a picture of "false clarity," yielding loadings that appeared "stronger," albeit misleadingly so. Using population models, Widaman's (1993) comparison of PCA to PAF found significant differences occurred with as many as 96 variables, presumably far more than a "reasonable" number of indicators. Fabrigar et al. (1999) found that PCA and PAF would lead to different interpretations for five of the18 data sets they examined. A hypothetical correlation matrix presented by Loehlin (1990) also demonstrated dramatically different interpretations if PCA were used in place of PAF.

In summary, there are cases where PCA will lead to interpretations no different than PAF. But there are also cases where the results will be different enough to affect interpretation of the data. Therefore, when preliminarily examining a construct that will be treated as an independent variable in line with the classical measurement model, PAF is a preferable factor extraction method (Floyd & Widaman, 1995). On the other hand, where the researcher is interested in developing a hypothesis about an emergent construct that will be conceptualized as a dependent variable driven by the indicators, PCA is an appropriate factor extraction method.

Eigenvalues Greater than One

E1 is the most commonly used criterion for deciding how many factors to retain. The most readily apparent weakness of E1 is its arbitrariness; a factor with an eigenvalue of 1.01 is regarded as indicative of a meaningful construct, while a factor with an eigenvalue of 0.99 is scrapped (Fabrigar, 1999). Zwick and Velicer (1982) tested E1 using samples drawn from data where the true number of factors was known. E1 consistently overestimated the number of factors, typically yielding one-third to one-fifth as many factors as indicators. A follow-up study by Zwick and Velicer (1986) also demonstrated that E1 overestimated the number of factors to retain. They concluded E1 should not be used. Widaman (1993) noted that the tendency of E1 to overestimate the number of factors, coupled with the tendency of PCA to become more inaccurate as the ratio of variables to factors decreases, exacerbates inaccuracy when the two procedures are applied in tandem.

The size of an eigenvalue is not directly related to the reliability of the composite scale (Cliff, 1988). Schönemann (1990) observed that any sample drawn from a population matrix equal to the identity matrix (in which all indicators are perfectly reliable and completely uncorrelated) would yield at least one factor with an eigenvalue greater than one, due to sampling error.

No consensus has emerged on universally applicable rules for deciding how many factors to retain. However, there does appear to be a consensus that of all criteria proposed, the E1 rule is among the very worst (Fabrigar et al., 1999; Gorsch, 1997; Hakistan et al., 1982; McCroskey & Young, 1979; Zwick & Velicer, 1982, 1986). The judgment of the researcher based on underlying theory might be the best method for making this preliminary determination.

Varimax Rotation

The goal of factor rotation is to make the factors more interpretable. The two main types of rotation are orthogonal and oblique. Orthogonal rotation imposes a condition on the data: that all factors must be uncorrelated. Oblique rotation does not impose this condition. Varimax rotation is orthogonal.

If factors are truly orthogonal, i.e., uncorrelated, then an oblique rotation will reveal this. However, the reverse is not true. Orthogonal rotation will misleadingly mask genuine correlations among the factors. Fabrigar et al. (1999) noted this is troublesome in psychology, where one would expect many constructs to be correlated in fact. Mass communication constructs also frequently correlate with one another. There is no reason to impose an orthogonal solution upon the data set. Rather, using oblique rotation will yield an orthogonal solution where the data support it, but also reveal the existence and extent of factor correlations where consistent with the data.

Confirmatory Factor Analysis

In contrast to exploratory factor analysis, confirmatory factor analysis actually tests the assumption of unidimensionality implicated by the classical measurement model. Hunter and Gerbing (1982) conceptualized confirmatory factor analysis as consisting of three basic steps: (1) confirm the face validity of each indicator; (2) estimate the factor loading of each indicator on the construct; and (3) test the measurement model for goodness of fit. *Face validity* (content validity) concerns whether there is a logical connection between each indicator and the construct. Assuming all indicators have face validity, the next step is to estimate the factor loading of each indicator on the construct. Because the construct is not directly observed, the "true score" of the construct must be inferred from the covariances of the indicators. A software program for these computations is available (Hamilton & Hunter, 1988). Structural equation modeling can also be used for confirmatory factor analysis (for a review see Kline, 1998).

Next, the model's goodness of fit must be tested. The first aspect of the goodness of fit test is to assess the *internal consistency* of the measure. The classical measurement model theorizes that each indicator is correlated with every other indicator only to the extent they each share common variance with the construct. Accordingly, if each indicator measures the same construct and only that construct, the correlations between the indicators will satisfy a product rule for internal consistency. For example, in the model set forth in Figure 1, the correlation between "Tells the Whole Story" and "Can be Trusted" should equal, within sampling error, the product of the factor loading of each on the newspaper credibility construct. This should hold true for every possible pair of inter-indicator correlations as well. If it does, the data support the internal consistency of the model. If not, it may be that one or more indicators should be dropped from the measure (such indicators might be grouped with others to form a measure of a related construct). If dropping indicator(s) still does not yield an internally consistent result, then new indicators might need to be devised. Alternatively, it is also possible that the unidimensional construct hypothesized by the model does not correspond to a phenomenon that exists in reality.

In some instances the researcher will work with more than one internally consistent measure. In such cases, these internally consistent measures need to be tested against one another. When indicators form unidimensional measures, they should exhibit similar patterns of correlations with indicators in other unidimensional measures. Hunter and Gerbing (1982) refer to this property as *parallelism*. Figure 3 demonstrates a simple path model for testing parallelism between measures.

INSERT FIGURE 3 ABOUT HERE

Again, a product rule is applied. The model hypothesizes that the correlation between indicators of two distinct unidimensional measures should be equal to the product of each indicator's factor loading on its construct and the correlation between the two constructs. Using the model in Figure 3 as an example, the correlation between indicators A_1 and B_1 should be equal to the product of the factor loading of A_1 on Construct A, the factor loading of B_1 on Construct B, and the correlation between Construct A and Construct B, within sampling error. This should also hold true for every other possible pair (consisting of one indicator from each construct) of indicators. This test of parallelism can be applied regardless of the magnitude or direction of the correlation between the constructs.

If the measure passes this test, then the data indicate the model is consistent with the parallelism theorem. If not, it may be that one or more indicators should be dropped from the measure (and perhaps used as part of a measure of a related construct). If dropping indicator(s) still does not yield consistency with the parallelism theorem, then new indicators might need to be devised, or it may be that one or both of the unidimensional constructs hypothesized by the model does not exist in reality.

Of course, a single confirmatory factor analysis does not establish the construct validity of a measure. Rather, a confirmatory factor analysis verifies whether data gathered in a particular study are consistent with the measurement model. Measures may withstand scrutiny in one population or setting, but fail to do so in others. In addition, it is also possible that a measure might withstand scrutiny when studied in the context of some variables, but then fail when studied in conjunction with some other variable. This might imply that the conceptualization of the construct requires refinement to account for this contextual difference. Hence, confirmatory factor analysis does not treat the construct validation process as having ended with the publication of a measure. Instead, the process continues as the measure is used in practice (Ozer, 1999).

Summary

In short, our review of mass communication journals suggests that studies rarely test for unidimensionality in connection with the development/assessment of new measures. The most commonly used techniques, Cronbach's alpha and exploratory factor analysis, do not examine unidimensionality. Failure to test the unidimensionally of a classical measure leaves the interpretability of the measure in doubt. In order to fully assess the construct validity of a measure, it is necessary to specify and test a measurement model. The classical measurement model hypothesizes that a single latent construct acts as an independent variable that is the sole antecedent cause (aside from error) of each indicator, i.e., this model hypothesizes unidimensionality, but mass communication studies routinely fail to test this implicit hypothesis.

Exploratory factor analysis is a preliminary technique that can be useful in the development of measures, but it does not test the goodness of fit of measurement model. Rote application of factor extraction techniques without due consideration of the underlying measurement model is ill-advised. Specifically, principal components analysis should not be used to develop a classical measure. Psychometric research suggests the eigenvalue-greater-than-one rule and varimax rotation are contraindicated.

If becoming a respected theorist is the highest achievement attainable by a mass communication scholar, then surely a special level of distinction must be reserved for those who elegantly model and rigorously test measures of latent constructs. Such measures hold promise to spawn countless conceptual theories.

Notes

1. Item covariances, therefore, cannot be used to judge the validity of a formative measurement model. Rather, it is necessary to expand the model and examine other variables that are effects of the latent constructs.

2. Kaiser first proposed this three-step EFA procedure (he previously invented varimax rotation itself) in his doctoral dissertation in 1956 (Kaiser, 1960). In 1964, Kaiser's colleague, Chester Harris, dubbed this three-step EFA procedure: "Little Jiffy," and so it came to be known among the leading psychometricians of their day (Kaiser, 1970). In 1974, Kaiser entirely overhauled the Little Jiffy procedure. Kaiser dropped PCA in favor of image analysis (another factor extraction method), and replaced varimax rotation with a modified version of quartimax rotation (Kaiser, 1974). Nevertheless, more than a quarter century after the inventor of Little Jiffy essentially discarded it, many mass communication scholars continue to publish studies employing the original Little Jiffy procedure, with no apparent regard for the measurement model that underpins their inquiry.

References

Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53,* 605-634.

Bollen, K.A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110,* 305-314.

Bollen, K.A., & Ting, K. (2000). A tetrad test for causal indicators. *Psychological Methods,* 5, 3-22.

Borgatta, E.F., Kercher, K., & Stull, D.E. (1986). A cautionary note on the use of principal components analysis. *Sociological Methods & Research*, *15,* 168-180.

Cliff, N. (1988). The Eigenvalues-greater-than-one-rule and the reliability of components. *Psychological Bulletin, 103,* 276-279.

Edwards, J.R., & Bagozzi, R.P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5,* 155-174.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272-299.

Floyd, F. J., & Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7,* 286-299.

Gorsuch, R.L. (1990) Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research, 25,* 33-39.

Gorsuch, R.L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality* Assessment, 68, 532-560.

Green, S.B., Lissitz, R.W., & Muliak, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Assessment*, *37,* 827- 838.

Hakstian, A.R., Rogers, W.T., & Cattell, R.B. (1982). The behavior of number-of-factors rules with simulated data. *Multivariate Behavioral Research, 17,* 193-219.

Hamilton M.A., & Hunter, J.E. (1988).*Confirmatory factor analysis.* Copies of the software may be obtained by contacting Dr. Hamilton, University of Connecticut, Department of Communication, Storrs, CT.

Hattie, J. (1985). Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9,* 139-164.

Hogan, R., & Nicholson, R.A. (1988). The meaning of personality test scores. *American Psychologist, 43*, 621-626.

Hunter, J.E., & Gerbing, D.W. (1982). Unidimensional measurement, second order factor analysis, and causal models. In B.M. Straw & L.L. Cummings (Eds.), *Research in organizational behavior* (Vol. 4) (pp. 267-320). Greenwich, CT: JAI Press.

Kaiser, H.F. (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Kaiser, H.F. (1970). A second generation Little Jiffy. *Psychometrika*, *35,* 401-415

Kaiser, H.F. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, *34,* 111-117.

Kline, R.R. (1998). *Principles and practice of structural equation modeling.* New York: The Guilford Press.

Loehlin, J.C. (1990). Component analysis versus common factor analysis: A case of disputed authorship. *Multivariate Behavioral Research, 25*, 29-31.

Loehlin, J.C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

McCroskey, J.C., & Young, T.J. (1979). The use and abuse of factor analysis in communication research. *Human Communication Research, 5*, 375-382

McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.

Meyer, P. (1988). Defining and measuring credibility of newspapers: Developing an index. *Journalism Quarterly, 65,* 567-574.

Ozer, D.J. (1999). Four principles for personality assessment. In L.A. Pervin & O.P. John, (Eds.), *Handbook of personality: Theory and research* (pp. 671- 686). New York: The Guilford Press.

Reise, S.P., Waller, N.G., & Comrey, A.L. (2000). Factor analysis and scale revision. *Psychological Assessment, 12,* 287-297.

Schöenemann, P.H. (1990). Facts, fictions, and common sense about factors and components. *Multivariate Behavioral Research, 25,* 47-51.

Snook S. C., & Gorsuch, R.L (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin, 106,* 148-154.

Velicer, W.F., & Jackson, D.N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research, 25,* 1-28.

West, M.D. (1994). Validating a scale for the measurement of credibility: A covariance structure modeling approach. *Journalism Quarterly, 71,* 159-168.

Widaman, K.F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research, 28,* 263-311.

Zwick, W.R., & Velicer, W.F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17,* 253-269.

Zwick, W.R., & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432-442.

Table 1

Percentage of Studies Using Various Methods to Test Measures of Mass Communication Constructs

Cronbach's Alpha Only	52
Exploratory Factor Analysis (w/o Confirmatory Factor Analysis)	44
Confirmatory Factor Analysis (w/ or w/o Exploratory Factor Analysis)	4

Note. N =124. Includes all studies that introduced a new measure of a mass communication construct and were published in *Mass Communication & Society* from 1998 to Spring 2002, *Journalism & Mass Communication Quarterly* from 1992 to Spring 2002, or *Newspaper Research Journal* from 1992 to Spring/Summer 2002.

Table 2

Percentage of Exploratory Factor Analytic Studies Using Various Techniques to Develop New Measures of Mass Communication Constructs

Factor Extraction Method	
Principal Components Analysis	58
Factor Extraction Method Not Reported	42
Decision Rule for Factor Retention	
Eigenvalue Greater than One	32
Decision Rule for Factor Retention Not Reported	68
Factor Rotation Method	
Varimax	59
Oblique	11
Both Varimax and Oblique	4
None	2
Rotation Method Not Reported	24

Note. N = 54. Includes all studies that used exploratory factor analysis to develop a new measure of a mass communication construct and were published in *Mass Communication & Society* from 1998 to Spring 2002, *Journalism & Mass Communication Quarterly* from 1992 to Spring 2002, or *Newspaper Research Journal* from 1992 to Spring/Summer 2002.

Figure 1

Newspaper Credibility Measurement Model, Adapted from West (1994).

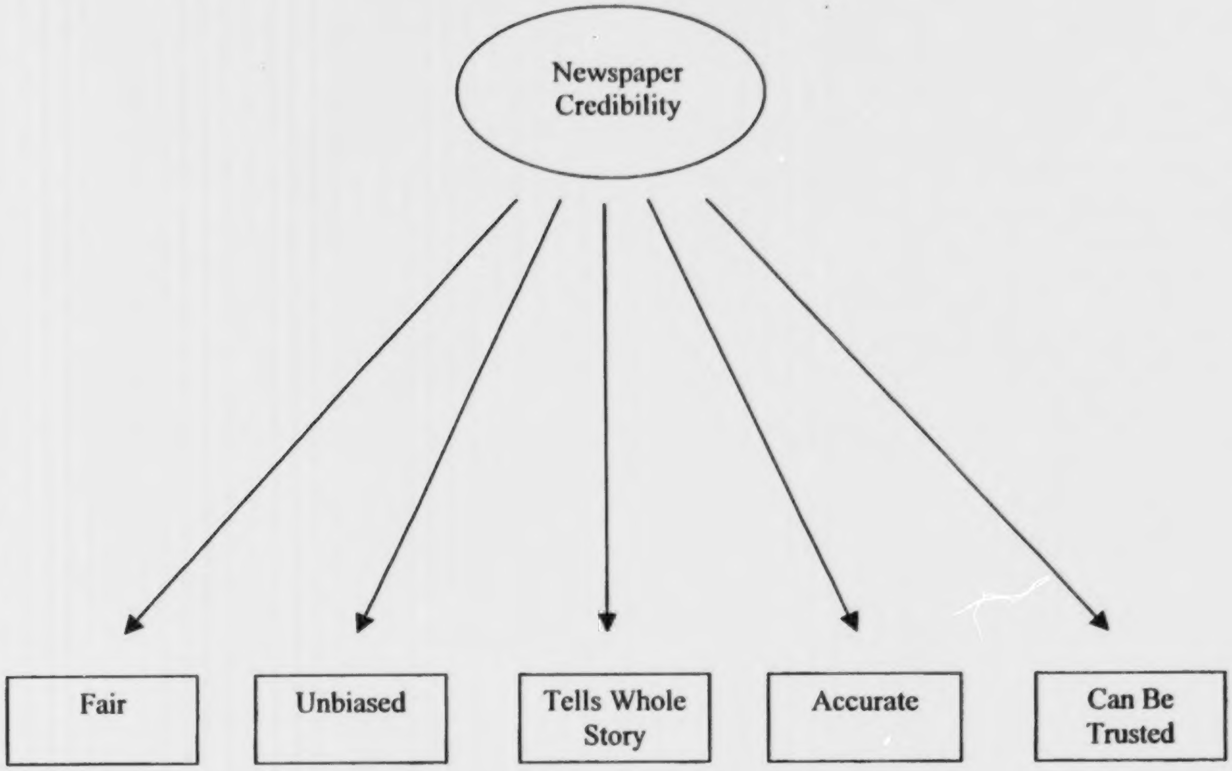

Newspaper Credibility
Fair
Unbiased
Tells Whole Story
Accurate
Can Be Trusted

Figure 2. *Violent Media Exposure Measurement Model.*

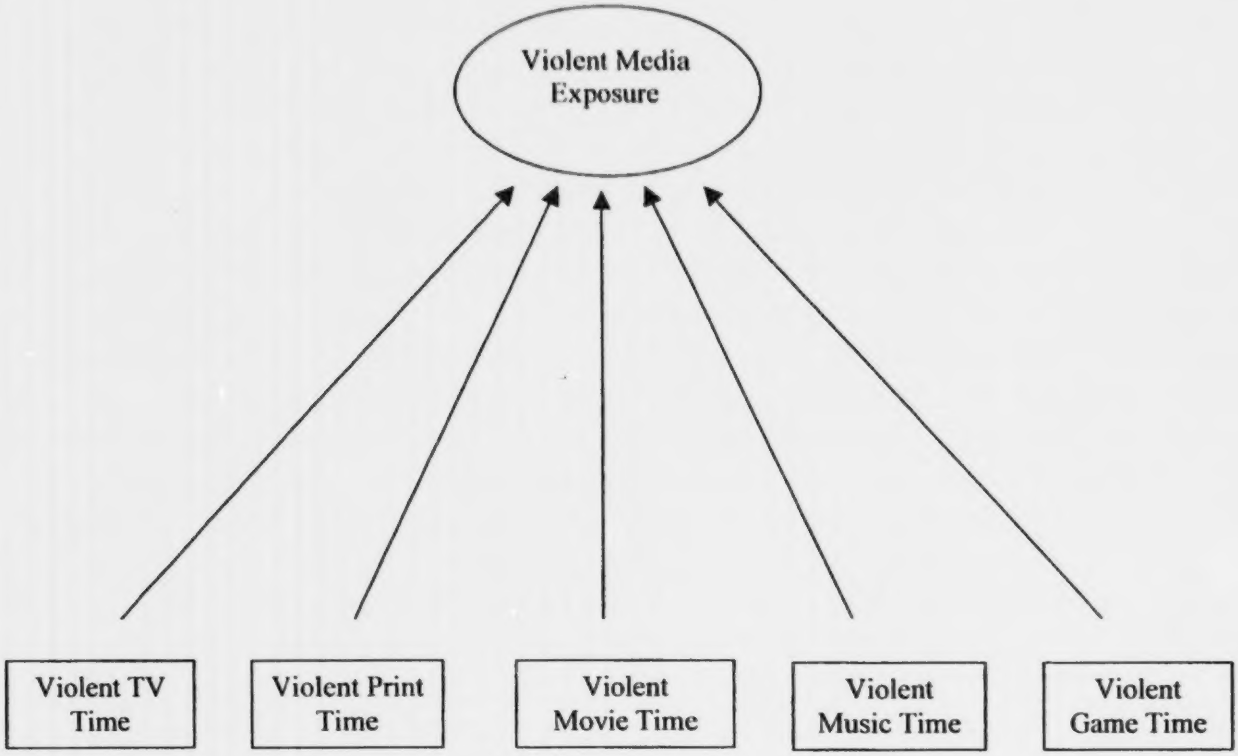
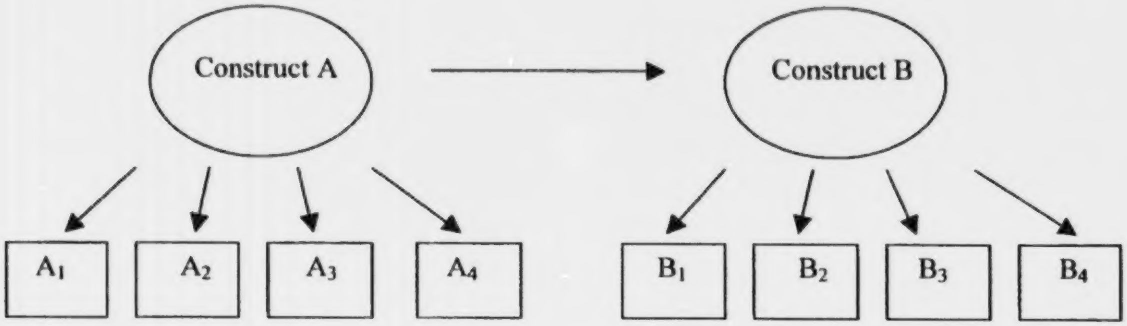

Violent Media Exposure
Violent TV Time
Violent Print Time
Violent Movie Time
Violent Music Time
Violent Game Time

Figure 3

Path Model for Testing Parallelism


Construct A
Construct B
A_1
A_2
A_3
A_4
B_1
B_2
B_3
B_4

Is Herpes Entertaining? An application of Entertainment-Education to text information processing concerning STDs among adolescents

by
Donna Rouner
C233 Clark
Colorado State University
Fort Collins, Colorado 80523, USA
9704915556
donna.rouner@colostate.edu

and

Ralf Kracke-Berndorff
227 S. Whitcomb Street
Fort Collins, Colorado 80521, USA
(970) 407-3171
ralf@ametrika.com

Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication, Communication Theory and Methodology Division, August, 2003.

Donna Rouner is a professor in the Department of Journalism and Technical Communication at Colorado State University, where Ralf Kracke-Berndorff is a Master of Science graduate.

Is Herpes Entertaining? An application of Entertainment-Education to text information processing concerning STDs among adolescents

This study examined the impact of entertainment-education strategies on audience's (N=137) information processing regarding sexual health. The hypothesis, that higher involved audience members exposed to a statistical message would show a higher message evaluation than those exposed to an anecdotal message, was partially supported. Also, "framing" the same anecdotal message as either intended for promotion or entertainment purposes, controlling for involvement, found higher message evaluation by respondents exposed to the entertainment message.

Is Herpes Entertaining? An application of Entertainment-Education to text information processing concerning STDs among adolescents

As of today, the only truly effective tool available to combat AIDS and certain other sexually transmitted diseases (STDs) consists of prevention, since effective pharmacological cures and vaccines have yet to be developed. Communication scientists could therefore play an essential role in alleviating the problem, especially as far as campaigns tailored for adolescents and young adults are concerned. This study will try to evaluate the effectiveness of different communication forms and styles on adolescents' beliefs and attitudes regarding sexual health. This study combines information-processing and entertainment-education (EE) approaches.

This study relied on fundamental theories of information processing, in particular in the realm of persuasion. Ajzen and Fishbein's (Ajzen, 1991; Fishbein & Ajzen, 1975) research in persuasion focuses on the question of how beliefs and attitudes actually translate into (or overlap with) appropriate behaviors. Their Theory of Reasoned Action states that a person's intention to perform a given behavior is a function of the person's attitude toward performing that behavior, the person's perception of the norms governing that behavior and the individual's motivation to comply with those norms. The researchers used open-ended procedures to explore audiences' beliefs toward specific behaviors. They also focused on beliefs regarding specific areas of the audience members' lives (behaviors, topics, social norms, etc.). Because beliefs have proved to predict behavior in a fairly reliable way, this variable constitutes a very practical tool in persuasion research.

Theoretical perspectives from cognitive response theory, such as the

Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986), stress the importance of a more active role of the receiver, especially if the receiver chooses to engage in a high thought elaboration while processing information through the so-called "central route." The authors suggest that a high-credibility source will lead to attitude change in one situation but not in another if attitude change is taking place through the peripheral route in one situation and the central route in the other. In other words, when the recipient of a message is motivated to process the message with some care, stronger arguments are influential (argument quality); when such motivation is absent, little attention is paid to the arguments in the message (argument quality matters little), and the peripheral route predominates. Hence, a person's involvement and the credibility of the source are key variables in predicting the impact of a message intended for persuasion.

Review of entertainment-education literature

While the ELM focuses on the processing of overtly persuasive information, other scholars have stressed the impact of entertaining (or fictional) information carrying "concealed" persuasive messages on the audience's beliefs and behaviors. Although praxis-oriented, this entertainment-education research borrows from a multitude of theoretical frameworks. EE tries to package prosocial messages into entertainment formats such as fictional prose, soap operas, music, radio shows, theater, etc. Although most of the entertainment-education projects have been conducted in the so-called "developing world," their implications help redefine new persuasion techniques in any kind of environment.

Brown and Singhal (1999) noted that the mass communication strategy of entertainment-education grew out of the recognition of two undesirable trends in contemporary mass media programming: the entertainment-degradation present in shows designed to shock the audience, such as sex, violence and voyeurism, and more purely educational programs that cannot maintain a sufficient number of viewers and are not commercially viable. The first trend is represented, for example, by Howard Stern's "shock radio" program, by television talk shows *à la* Jerry Springer, or by certain tabloid newspapers; the second by the failures to produce popular children's education programs for commercial networks (Brown and Singhal, 1999). The relatively new EE approach tries to overcome these two trends by blending their positive aspects; increasingly, producers try to create programs that are entertaining, educational, socially responsible and commercially profitable at the same time.

Entertainment television programs have influenced increasingly large audiences globally during the 1990s; especially in "developing" countries where most of the world's population resides, this growth is most evident. The People's Republic of China counts 700 million television viewers, India more than 90 million. One of the most popular television genres in the world is the soap opera. Soap operas, or *telenovelas*, are the most popular type of television programs in Latin America and are rapidly growing in popularity in Asia (Brown 1992a, 1992b).

According to Piotrow (1994), key elements of entertainment make it especially useful for disseminating pro-social messages that provide education and advance development goals. Entertainment is perennial, pervasive (expanding in all countries), popular (across cultures), personal (can present educational material in a more personal way, without the embarrassment involved in addressing sensitive issues interpersonally), pleasurable (release or escape), persuasive (encouraging the audience to adopt pro-social attitudes and behaviors), passionate (stirring up emotions), profitable, and practical. The use of entertainment-education strategies in television is credited to a creative writer-producer-director of television in Mexico, Miguel Sabido. He was inspired by the unintended educational effects of a successful Peruvian soap opera, "Simplemente María," and pioneered the EE strategy in television (Singhal, Obregon, & Rogers, 1994). Sabido based his strategy on a message design framework that incorporates several theories, including social learning and social marketing theories.

Bandura, in his social learning, and later, social cognitive theory (1977, 1994), argues that humans learn social behaviors by observing role models. Social networks, an important component of Rogers' (1995) Diffusion of Innovations theory, are essential for the spreading of newly acquired behaviors. Bandura also emphasized the self-confidence of an individual to be able to perform new skills will impact the success in changing a given behavior. This perceived self-efficacy affects every phase of personal change. Information campaigns using social modeling are therefore aimed at building the receiver's self-efficacy, and not only to convey knowledge and rules of behavior (Bandura, 1994). Particularly in health communication, this approach proved to be more effective than the mere transmission of factual information, fear arousal, and change in risk perception (Meyerowitz & Chaiken 1987).

Sabido's entertainment-education soap operas utilized this approach, with prosocial behaviors rewarded, antisocial behaviors in relationship to the topics of the soap operas (Singhal & Rogers, 1999). Although entertainment-education proved to be effective in conveying prosocial messages to the audience (ultimately leading to behavior change), and some research has been done in monitoring its effects, a comprehensive theoretical framework has yet to be developed that explains the processing of the stories among audience members. Some more recent concepts may help illuminate the EE phenomenon from an information-processing point of view. Green and Brock (2000) emphasize the importance of *transportation* in the persuasiveness of narratives. Transportation theory assumes that the reader (viewer, listener) of narratives abandons his/her own world, and is transported into the story. When the "traveler" returns to his/her world of origin, he/she is somewhat changed by the journey (Gerrig, 1993). Persuasion scholars, who concentrated their research mostly on cognitive elaboration, have largely ignored transportation. Green and Brock state that transportation is a convergent process, whereas elaboration might be conceived a divergent process. Rather than having a single focus (e.g., the narrative), a person engaged in elaboration might be accessing his or her own opinion, previous knowledge, or other thoughts and experiences in order to evaluate the message at hand. Under high elaboration, connections are established to an individual's other schemata and experiences. In contrast, under high transportation, the individual may be distanced temporarily from current and previous schemata and experiences, a circumstance likely to favor an effective persuasion.

Brosius and Bathelt (1994) pointed out the utility of exemplars, or illustrative individual cases in persuasive communication. They found a stronger impact of exemplars on the perception of a problem than of general statements about the importance of a problem (base-rate information). Although their study focused on news reporting on an issue, it emphasizes the importance of exemplars on *perceived persuasiveness* and beliefs.

Subsequent research in persuasion (cf., Reinard, 1988) suggests that evidence, when reasonably well referenced, does enhance the persuasive effect of a message. Two frequently used types of evidence are the statistical and the anecdotal (corresponding to the above-mentioned base-line information, and to the exemplars, respectively). In recent research, Slater and Rouner (1996) distinguished between value-affirmative and value-protective message processing. They found that statistical evidence was useful to positively reinforce the beliefs of the "already converted" (value-affirmative), whereas it had a lesser (or even an inverse) effect on informants with an incongruent belief system (incongruent toward the evidence contained in the message). Anecdotes or exemplars, on the other hand, had an effect primarily on the value-protective recipients (incongruent beliefs). Unlike Petty and Cacioppo's (1986) research, their study did not compare attitudes at two times, but measured participants' perceived persuasiveness of the message.

Block and Keller (1997) studied the effects of vividness on the *persuasiveness* of health messages. The researchers' approach was similar to Slater and Rouner's (1996), measuring informants' perception of persuasiveness.

In a recent study, Slater and Rouner (2002), attempted to put entertainment-education in the context of Petty and Cacioppo's ELM. While the latter focused on explicitly persuasive messages, it did not consider narrative, which involves absorption and transportation of the recipient. They found a clean distinction between central and peripheral processes is no longer discernable in the case of narrative processing.

The present study examines the impact of a persuasion message using a derivative EE approach, comparing the differences between narrative (entertaining) and statistical (informative) messages, manipulating types of message and message sources, whether entertainment or persuasion, to determine their impact of message evaluation. We use key concepts from the reviewed theories to investigate information processing of an EE message: beliefs, involvement, self-efficacy and source credibility.

As already mentioned, Ajzen and Fishbein approached their audiences with open-ended questions in an effort to discover their beliefs directed toward specific behaviors, topics, or social norms. Not only do beliefs predict subsequent behaviors, but they also influence cognitive processes in persuasion.

According to ELM, depending on the person's motivation, interest or involvement in a certain topic, the message will be either processed through the central, or the peripheral route. In this study, involvement is used as a control variable, further explaining audience's differences in processing. If differences are found across statistical and anecdotal texts, the level of involvement may help explain the differences. High involvement predicts central processing, in which argument strength seems to be a predominant factor, whereas when involvement is low, argument quality does not seem to matter, and the message is processed through the peripheral route. According to Slater's (1997) "extended ELM," and Slater and Rouner (2002), who applied that model to entertainment-education, such a clear distinction between central and peripheral routes seems to vanish.

Self-efficacy has been used as a trait-like concept that characterizes individuals' general tendency to feel powerful in impacting their own lives. More recent research

(Rosenthal, Moore, & Flynn, 1991) has conceptualized self-efficacy as more situationally determined. Thus a measurement of self-efficacy specifically focused on preventative health behavior in sexual matters is more useful to this study. In their study on adolescents' sexual risk-taking behavior, Rosenthal *et al.* adopted a sexual self-efficacy scale (SSE) from Libman, Rothenberg, Fichten, and Amsel (1985) to measure adolescents' self-efficacy.

The concept of source credibility was introduced more than 50 years ago by Hovland and Weiss (1951). Results showed that the informants discounted information from untrustworthy sources. In the context of the Elaboration Likelihood Model, source credibility is considered "peripheral" to the message itself, and assumes a particular importance for audience members processing the information in a state of low involvement. Involvement is not as important in predicting central and peripheral processing with entertainment messages, a message type for which the importance of source credibility is relatively unknown. In any case, source credibility constitutes a fundamental variable in sexual health information processing, especially if one considers sexually transmitted diseases that are potentially life threatening. This study uses two basic types of stimuli, consisting of one statistical and one anecdotal message, an important distinction Brosius and Bathelt (1994) introduced in persuasion research.

"Framing" of the message

One way to study message effects in an entertainment-education context is to examine the difference between how people process entertainment-versus persuasive messages, and one way to do this is to cue audience members' schemata relative to a promotional (educational) or an entertainment script. This cueing of audience members' schemata--while leaving the actual messages identical--may lead to interesting differences in information processing. So-called "unitization" research is based upon the idea that only certain key "frames" or "chunks" of incoming information are encoded and stored to enable a person to reconstruct the information later (Reeves, Chaffee, & Tims, 1982). An operational schema, especially one representing processing goals, will determine how a stream of incoming information is unitized or encoded (Cohen, 1981; Reeves et al., 1982). Hence, introducing, or "framing," the same message either as intended for promotional, or as intended for entertainment purposes, may effect people's perception of the message itself.

Schema theory is also useful in studying the self-concepts of individuals who are targeted by EE efforts. Self-schemas, for example, have been shown to moderate the relationship between intentions and actual behavior (Sheeran & Orbell, 2000). This area of social cognitive theory may be enhanced with a more affective approach, a look at the development of future self-concept, including constructs like "future self," "possible self," or the "idealized self" (Higgins, 1987; Kato & Markus, 1993). These schemas about the self would presumably reveal perceptions about whether one held sufficient knowledge about topics covered by EE messages, how adequately one used information on these topics, and possibly one's inclination toward future behavior change. Late adolescents, from 15-20 (Millstein, Petersen & Nightingale, 1993; Seiffge-Krenke, 1998), who are in the process of forming a more permanent and unique identity, are in an important life cycle change where strong media and interpersonal communication effects might be likely. These might range from deleterious media content as well as prosocial preventive health campaigns.

This study will not examine the effect of the two message types on the audience's knowledge gain, or belief- and behavior change. It rather measures the perceived "message usefulness" which will ultimately impact audience's proneness to change beliefs and behaviors in the direction of the message content (see "perceived persuasiveness,"Block & Keller, 1997; Brosius & Bathelt, 1994; Slater & Rouner, 1996). In other words, the audience is asked to evaluate the message according to its persuasiveness, entertainment-value, plausibility, etc.

Given the nascent stage of looking at entertainment-education from the viewpoint of audience's information processing, instead of a more sender-oriented perspective, it seems at this point appropriate to replicate the findings of the above-mentioned research showing the advantages of specific genres for audience's need (Brosius & Bathelt, 1994; Slater & Rouner, 1996). Specifically, Slater and Rouner's (1996) findings indicated that people whose beliefs were consistent with the message content were more impacted by statistical information, whereas those inconsistent with the message rated the anecdotal information higher. We therefore advance the following hypotheses:

H1: Audience members exposed to the statistical message and who are more involved in sexual health issues, will show a higher message evaluation than audience members exposed to the anecdotal message. Conversely, audience members exposed to the anecdotal message, and are less involved in sexual health issues, will show higher message evaluation than audience members exposed to the statistical message.

Investigating the impact of an entertainment-education soap opera in India, Singhal and Rogers (1989) found that overwhelming emphasis on educational messages decreased the perceived attractiveness of the shows. This is in accordance with Sabido's theory of tones (Brown, 1992a). Exaggerating educational components negatively impacted the entertainment value of the episode. Following the above mentioned unitization approach (Cohen, 1981; Reeves et al., 1982), it would be expected that simply "framing" the same message as either entertaining or promotional will effect audience members' evaluation of the message itself. Hence:

H2a: When exposed to the same anecdotal messages, framed either as intended for health promotion purposes, or as intended for entertainment purposes, audience members exposed to the message framed as intended for entertainment will evaluate the message higher than audience members perceiving the message as promotional.

As already mentioned, according to Bandura's (1994) social cognitive theory, people will evaluate message differently depending on their self-efficacy. Likewise, in their Elaboration Likelihood Model (ELM), Petty and Cacioppo (1986) stressed the importance of issue-involvement in audience's information processing. Finally, Ajzen (1991) and Fishbein and Ajzen (1975) emphasize the effects of beliefs on behavior change. We will therefore advance the following hypothesis:

H2b: Hypothesis 2a will hold up controlling for involvement, self-efficacy, and beliefs.

The identification of the observer with the model is of central importance in vicarious learning (Bandura, 1994); people identify with other's character traits, beliefs and actions, and observe the consequences of their behaviors. The closer the observer perceives him or herself to a character, the more he/she will be affected by the learning experience. This homophily is an important part of diffusion of innovation (Roger, 1995), as well. Accordingly, it seems especially relevant to identify with a same-sex character in a story about a sexually transmitted disease. We therefore advance the following hypothesis:

H3: Audience members who read a story featuring a protagonist of their same sex will rate the message higher. In other words, a gender-homophily is expected for audience members exposed to the narrative.

It would be interesting to extend H1 with the variable of source credibility, given the importance of source credibility in persuasion research, as well as in the Elaboration Likelihood Model. Slater and Rouner (1997) found that message type impacts source credibility. We therefore pose the following research question:

R1: Does the message type (anecdotal versus statistical information) impact perceived source credibility?

Likewise,

R2: Does the "framing" of the message impact perceived source credibility?

Gender differences have been found in the evaluation and processing of advertising messages (Rouner, Slater & Domenech-Rodríguez, 2003). We would therefore like to pose the following research question:

R3: Does gender impact message evaluation?

Finally, we propose a qualitative look at adolescents' self-concepts, present and future, in order to reveal patterns of how they perceive themselves as decision-makers about their sexual health.

Methods

Informants

The study consisted of a between-informants experiment conducted on first-year students enrolled in an advertising class at a large-size Western university. The 137 participants were randomly exposed to one of six experimental conditions, consisting of different message types (anecdotal or statistical) and differing types of sources (entertainment or persuasive), with the non-statistical, narrative messages presenting either a male or female protagonist with whom one might identify.

Procedures

During their routine class period, students in an introductory advertising class Who volunteered to participate were briefly informed about their ethical rights and then completed the experimental instrument they were randomly assigned.

Instrument

The questionnaire opened with a set of open-ended questions on personal decision-making, in an attempt to measure self concept. Participants were asked to describe themselves in the role of a decision-maker regarding health decisions about sex. Subsequently, they were asked to evaluate their decision-making capabilities in the future ("a few years from now"), to assess future self concept.

The third open-ended question was intended to measure the raw knowledge of the participants regarding sexually transmitted diseases; the students were asked to list the STDs they were familiar with, and to briefly describe their respective symptoms.

The instrument provided a list of 25 belief pairs obtained in a pretest, and organized in a seven point semantic differential: The pretest, using 37 adolescents similar in age conducted at a Western high school, yielded the belief items for the main experiment. In the pretest, Fishbein & Ajzen-like (1975) open-ended questions regarding beliefs about sex provided a range of beliefs out of which the 25 most common provided the basis for a 7-point semantic differential scale. Sample items include: "passionate," "exciting," "embarrassing," "romantic," "emotional," "tiring," "OK before marriage," "painful," etc, with respective opposites.

The instrument presented a list of ten statements regarding participants' involvement in (sexual) health issues. Students were asked to choose from a five-point Likert scale (from 5 = "strongly agree," to 1 = "strongly disagree"). Sample items include: "I pay close attention to health information about sex-related issues," "My friends come to me for up-to-date facts about health issues," or "I'm confident that I make sound decisions about my health."

To measure self-efficacy in sexual health decision-making; students were confronted with a list of 14 behaviors (Rosenthal, Moore, & Flynn, 1991) and for each asked to choose from a five-point Likert scale (5 = "absolutely certain," 1 = "very uncertain"). Sample items include "Discuss using condoms and/or other contraceptives with a potential partner," "Be able to buy condoms/contraceptives," "Put a condom on a penis," and "Discuss sex-related issues with a family member."

To create the message stimuli (Appendix), data from the Centers for Disease Control (CDC) and health related web sites, applicable to males and females, were adapted to a statistical text about *Herpes Simplex II*. This set the baseline for a statistical text. Identical anecdotes with male (David) and female (Cynthia) protagonists were derived from the statistical text, using the same information. All three conditions had the same length (498 words) and were equal in readability scores (Severin & Tankard, 1997).

Herpes was chosen, and in effect favored over AIDS/HIV, because it constitutes a common disease that affects young people across social and racial strata. It is a serious disease, which seems to be neglected and somewhat overshadowed by the predominant presence of information on AIDS/HIV in the media. By choosing herpes we tried to avoid redundancy, boredom and excessive fear arousal in the audience, while still creating a realistic, generalized, and serious scenario.

An introduction before each message provided a statement concerning the intended use of the message: one introduced the message as a "basis for the creation of a health information brochure," the other as "the basis for the creation of a dramatic television series." Thus the message stimuli comprised six types (statistical, anecdote male, and anecdote female times two).

Over 50 years of source credibility research have yielded two dominant dimension of source credibility, consisting of trust and expertise. Trust was considered more relevant for this study, given that the subject of the study concentrated on decision-making about sex. This study uses six items of the trust dimension ("believable," "trustworthy," "accurate," "biased (this item was later reversed)," "fair," and "honest"), and two items of the expertise dimension ("authoritative" and "knowledgeable"), assessed with a 5-point Likert-like scale.

A similar Likert-like scale adapted from Slater and Rouner's (1996) perceived persuasiveness scale was used to determine message evaluation. This scale was integrated with items related to narrative messages. Sample items include "believable," "convincing," "persuasive," "not useful (later reversed)," "realistic" and "tells a story."

Participants were asked to rate sources of information seeking concerning sexual issues, first according to personal preference, then trustworthiness. They ranked their top three choices of 16 items, including "family members," "magazines," "Internet health sites," "television news," "entertainment films," "health professionals."

Coding of the open-ended questions about self-concept was conducted by the two principal researchers, who derived categories from reviewing the open-ended data using the qualitative research techniques of determining common themes and placing comments into categories based on those these. Intercoder reliability was determined to be adequate (Cohen's kappa= .74), using a 15% subsample of the data.

The questionnaire included the following demographic questions: gender, ethnicity, sexual orientation, frequency of sexual intercourse, amount of children, and marital status. We also asked the students to indicate their number of past formal presentations on sexual health.

Results

Out of the 137 participants, 50% were male, 50% female, 94% Caucasian, 3.1% Hispanic, 1.6% Asian, 0.8% Native American, and 0.8% were Blacks, a demographic distribution that closely matches the demographics of this university. The participants reported the frequency of their sexual activity as follows: 9.1% "daily," 6.8% "4-5 times a week," 24.2% "1-3 times a week," 15.9% "2-3 times a month," 17.4% "once a month," and 26.5% "not active." Participants were asked to rate their most *preferred* and most *trusted* sources of information about sex; unfortunately, 33% of the respondents overlapped their ranking; out of the remaining sub-sample, the results were as follows: 29% listed health professionals as their most *preferred* source of information, 22% mentioned personal conversations with health professionals, 15% preferred friends, and 9% family members. The second most preferred source of information consisted of health professionals (20%), friends (20%), family members (8.7), personal conversations with health professionals (9%), printed information with health professionals (9%), and magazines (10%). Rank three consisted of friends (22%), family (14%), health professionals (13%), and personal conversations with health professionals (11%).

When participants were asked to rank their most *trusted* sources of information about sex, they ranked as first health professionals (39%), conversations with the health professional (26%), family members (8%), and friends (7%). Rank two consisted of health professionals (17%), friends (16%), printed information from health professionals (12%), family members (11%), and personal conversation with health professionals (10%). As their third most trusted source, participants mentioned friends (23%), personal conversations with health professionals (11%), and family members (14%).

Regarding the informants' ability to name STDs, 45% of the adolescents mentioned HIV/AIDS first, 21% noted it second and 23% noted it third. The only other STD with high salience was genital herpes, which was mentioned by 31% of the respondents first, 37% of the respondents second, and 13% of the respondents third. Two STDs for which the risks are highest among this age group are human papilloma virus (HPV) and chlamydia. Six percent of the respondents noted HPV on first mention, 6% on second mention, and 14% on third mention. As for chlamydia, 7% mentioned it first, 7% second and 10% third. On average, the informants could name about 2 STDs, but they could only accurately identify about 1½ symptoms.

Out of the 100 participants who answered the question: "How many formal presentations have you heard on today's topic (for example classroom presentations or workshops)?", the average number of presentations was 4.6 (from the remaining 37, 3 responded "many," 19 responded "a lot," and 16 did not respond).

For the self-efficacy measure, a confirmatory factor analysis using Varimax rotation yielded a five-item solution that accounted for 67.38% of the variance. The first dimension, which accounted for 24.46% of the variance, was used to create an index of self-efficacy. Eight items ("discuss using condoms with a partner," "ask a potential partner to wait if precautions are not available," "carry condoms around 'just in case'," "choose when and with whom to have sex," "discuss precautions with a doctor," "admit being inexperienced to your sexually experienced peers," "discuss sex-related issues with your partner") seemed to indicate people's comfort level in performing specific tasks efficaciously, but did not necessarily indicate a generalized comfort in all areas. The eight loaded items were used to create a summative index of self-efficacy, with a Cronbach's alpha of 0.83.

Pretests based on Fishbein and Ajzen's open-ended beliefs items yielded a list of 25 items, organized in a 7-point semantic differential. An exploratory factor analysis using Varimax rotation resulted in six dimensions and accounted for 62.29% of the variance. The first dimension was used to create a beliefs index, accounting for 22.70% of the variance. The 11 items loading strongest for this belief dimension characterize the sensationalized, more casual, light-hearted aspects of sex ("passionate," "exciting," "romantic," "not stressful," and "not embarrassing"). Items that were rated positively or negatively were recoded into the same direction. A higher number represented a more positive belief (e.g. "happy," "exciting," and "not stressful"). They were added to form a scale with a Cronbach alpha of 0.83.

Out of the ten Likert-like involvement items, a Varimax rotation yielded three dimensions that accounted for 64.38 of the variance. The first dimension accounted for 28.26% of the variance. Five items from the first dimension, loading greater than 0.48, were summed to form a scale (Cronbach's alpha=0.78). These items seem to represent the use of health related information sources for a sound decision making (e.g., "I pay close attention to health information about sex-related issues," "It's important where I get my health information," "My friends come to me for up-to-date facts about health issues," "I know exactly where to find what I need to know about health issues").

A Varimax rotation of the 18 message evaluation items yielded 5 dimensions, which accounted for 68.85% of the variance. Six items from the first dimension accounted for 22.08% of the variance; these items seem to correspond to an overall perceived message quality ("convincing," "persuasive," "well-written," "interesting," "important," and "beneficial"). This index had a Cronbach's alpha of 0.87.

Regarding source credibility, a confirmatory factor analysis with Varimax rotation yielded two dimensions; 7 of the 8 items loaded on the first dimension, accounting for 49.97% of the variance. These items were "believable," "trustworthy," "accurate," "fair," "authoritative," "honest," and "knowledgeable." An additive index of these items showed a Cronbach's alpha reliability of 0.88.

Manipulation Checks

The statistical and the anecdotal stimuli were found to significantly differ relative to a single-item measure of whether the respondent determined the text to be telling a story or to be of a statistical nature ($F = 99.44, p < .001$; $F = 15.18, p < .001$, respectively). In addition, the framing measure showed a similar relationship with a self-report item of how entertaining the text was (F = 6.49, $p < .05$).

Hypothesis testing

To test **H1**, which stated that audience members who are exposed to the statistical message and who are more involved in sexual health issues, will show a higher message evaluation than audience members who are exposed to the anecdotal message, and that, conversely, audience members who are exposed to the anecdotal message, and are less involved in sexual health issues, will show higher message evaluation than audience members exposed to the statistical message, the sample was divided into those reading statistical messages, and those reading anecdotal texts. Regressions were run within these subgroups, predicting message evaluation as a function of involvement in using informational sources in sexual health decision-making. For the statistical group, involvement predicted message evaluation (beta = 0.39, $p < .01$; Table 1). However, for the anecdotal group, this was not the case.

To test **H2**, which stated that, when exposed to the same anecdotal messages, framed either as intended for health promotion purposes, or as intended for entertainment purposes, audience members perceiving the message as entertaining will evaluate the message higher than audience members perceiving the message as promotional, and that this would hold up when controlling for involvement, self-efficacy and beliefs, we ran a t-test, which turned out to be not significant, although it pointed in the predicted direction. However, controlling for involvement, beliefs, and self-efficacy items, we found that, when controlling for involvement, this framing

Table 1

Summary of Regression Analysis for Involvement Predicting Message Evaluation within the Anecdote and Statistical Subgroups (N = 134)

	Variable	B	SE B	β
Anecdotal Group				
	Involvement	*-0.04*	*0.10*	*-.05*
Statistical Group				
	Involvement	*0.38***	*0.13*	*.39*

**$p < .01$

yielded a difference in message evaluation (F = 4.84, $p < .05$; Table 2). The mean for entertainment-framing was 3.97; the mean for promotional framing was 3.68.

Table 2

Analysis of Variance for Message Evaluation

Source	df	F	η	p
Involvement	*16*	*1.72*	*.93*	*.17*
Framing	*1*	*4.84*	*.51*	*.05**

*$p < .05$

When testing **H3**, which stated that audience members who read a story featuring a protagonist of their same sex rated the message higher, we did not find it significant; however, the results pointed in the hypothesized direction.

Research Questions

Research Question 1, which asked whether the message type (statistical versus anecdotal) influenced source credibility, proved positive. People exposed to the statistical message ranked source credibility higher then those reading an anecdote (statistical group mean=3.88; anecdotal group mean= 3.64; t = -2.13, $p < .05$).

Research Question 2, which asked whether the "framing" of the message type influenced source credibility, did not yield any significant results. Research Question 3, which asked whether gender impacted message evaluation, proved positive: we found that overall, females rated the messages higher than males (female mean = 4.04, male mean = 3.71, t = 2.84, $p < .01$).

The self concept analysis yielded some interesting patterns of present and future views about one's sexual health decision making and information use. Principally, the informants broke into three major groups, those who revealed they were not particularly good decision makers, those who felt they exerted some control now, but would be even more so in the future and those who were sexual virgins, demonstrating resolute control present and future. The following quotes will help portray these trends in the data.

Not a good decision maker

Right now, I'm pretty risky. I'm informed of diseases and aware of my actions/consequences, but I usually don't use condoms when having sex. The information and resources are there, I just choose not to use condoms for personal pleasure (for me and her). I think as I evolve as a more experienced player, I'll be able to gauge what types of girls to "strap one on for." My maturity level will be higher, therefore allowing me to think about wearing condoms, regardless of how dirty the girl is.

White male, quite knowledgeable about STDs

I'm not a good decision-maker about sex. I know little about that. I'm a foreign student from China. It's such a traditional country that the information about sex is limited. In US, people usually more open-minded than Chinese do. So, I think I can get more and a large variety of information about sex here. It's better because I will be in better control of decisions about sex.

Chinese female, could only mention AIDS

Depends if I am drunk. If sober, good decisions. If drunk, bad decisions. Hopefully I won't have any STDs in 5 years. Also, my decisions will be better.

White male, some knowledge

I am interested in ways of having safe sex. Birth control pills get to be too expensive and condoms are inconvenient.

White male, 3 STDs, but no symptoms

I am a bad decision maker about sex. I don't use protection. I will be better because I will be more informed and I won't be as desperate.

White male, minimal knowledge

Some control now--even more in the future

[I'm] a safe decision maker and feel in control of situations; more so than in the past. I will probably be more safe as I get older. I will continue to give p on the invincibility of youth.

No identifier

I think I am a pretty strong decision maker regarding health decisions about sex. [The] majority of my information about sex stems from my past--things like sex education] in middle and high school, parents and my doctor. I don't like condoms and I don't use them. I only engage in sexual activity with people I trust (know are clean). I take birth control pills. In a few years I'll be 22. I think I will be even more in control of my sexual decisions, I think there will be less temptation when I am that age and there is also a good chance I'll be in a committed monogamous relationship. Most of my information sources will probably be my doctor, school and research findings.

White female, fair knowledge of 3 STDs

I think I could be more selective with who I sleep with. My decisions are definitely altered by alcohol. I see myself living happily with no kids. I'm a big supporter of birth control.

White male, fair knowledge

I feel pretty educated in safe sex. I take birth control pills and use condoms. I feel pretty confident in my level of control. I plan on getting married around 25 so I will be on birth control still. I still want to use condoms because you never know what can happen.

White female, good knowledge

I feel that I'm well informed. I have a good idea of the risks associated with sex and how to prevent them. But, overall, I'm the type to get caught up in the moment and worry about risks later, bad I know, but honest. I can feel myself becoming more motivated as we speak about being safe. I don't see myself doing more "research," but more of using what I know.

White male, symptoms: HIV, The Clap; Symptoms: "No idea"

Pretty good in decision making except for the few instances where alcohol is involved. I am in control for the most part. I think that I will probably be the same.

White female, minimal knowledge, alcohol

Right now I am a much better decision maker about sex. Just recently I had a scare and realized how much importance and impact each decision had on my life. I'd like to think I will sustain from it unless in an intimate relationship, a long term one.

White female, mentioned 1 STD plus symptoms

I feel that I am very responsible in regards to my decisions about sex. I am in complete control of those decisions. In several years, my sexual activity will be much less than now (fewer partners). I will probably not feel as much need to learn about risks etc.

White female, no STDs mentioned

Virginity

I choose not to have it! I am sure I will look into sexual health decisions closer to marriage, but I already know how to get the information.

White female, named two STDs

I am a virgin and plan on staying one until marriage. When I become sexually active I plan on having safe sex, possibly using birth control, but I plan on only having one partner. If I happened to get pregnant before I was ready, I would have an abortion.

White female, some knowledge, virgin, yet pro-abortion

I have pretty good morals and decision making and at this time I have not had sex. One thing I fear is that I will when I'm drunk. Hopefully if I am not married, I will still be a virgin or at least, if not, I will have done it with someone I love.

White female, good knowledge, alcohol!

I choose to remain abstinent because of religious beliefs. I will have sex when I get married. In my decisions about sex I know I will not change. Sex is something married couples share.

White male, some knowledge

Some recurrent themes in the data revealed most respondents considered themselves knowledgeable and in control, and even predicted improvements for the future; alcohol was mentioned frequently. Many respondents admitted to not always make the right choices based on their knowledge (inconsistency). Some had concrete ideas about the future, for example, "…at 23 I will have 2 kids."

Discussion

Hypothesis 1, which essentially consisted of a replication of Slater and Rouner's (1996) findings on alcohol use, was confirmed only for the statistical part. The fact that a more issue-involved audience will use statistical evidence to support its views and values is fairly intuitive: people already somehow aware of an issue, and already leaning toward attitudes and behaviors advocated in the message, will be naturally more likely to rate a statistical message higher than audience members showing a low involvement. These low-involved people may be apathetic, or they might be selectively perceiving or respectively avoiding the message. For the highly involved audience members, statistical evidence provides reinforcement and legitimization of their views and opinions. Unfortunately, in this study we did not find evidence of the second half of the hypothesis, which would have provided a stronger contribution to the EE approach in the design of effective messages for a hard-to-reach audience.

This might have occurred due to the relatively weak stimulus consisting of the piece of narrative. Most entertainment-education messages employ longer and more engaging stimuli, which go beyond a 498-word long narrative, including other types of media such as radio shows, television soap operas, video-games, popular music, and fictional prose. The principal weakness of the message presented in the course of this study consists of the fact that its shortness probably did not allow a true identification of the audience with the portrayed character in the story; the characters might not be sufficiently developed. However, the use of two written pieces of information of the same length and readability scores not only seemed more feasible for the purpose of this study, but also avoided the confounds inherent in other types of media. For example, it would have been far more difficult to create two video clips containing the same amount of information: far too many unpredictable variables and confounds would have been introduced to allow a valid quantitative study. These variables include such complex stimuli as physical features of the actors, mimic, voice timber, and formal features such as camera angles, sound, music, focal length, or editing techniques. However, one might argue that the same narrative could have been presented in a video format by means of either a talking head reading the statistical message, or by an actor (male and female, respectively) who simply tells his/her story in front of the camera. This hybrid solution would indeed minimize the risk of stylistic confounds, since the anecdote is not acted out but simply spoken by the protagonist. However, this format would still allow little identification with the protagonist, since it does not take advantage of the cinematographic features of dramatic storytelling through the modeling of actors, nor would it facilitate the transportation of audience members.

However, Slater and Rouner (1996) had also used two written messages; the question remains why this study was not able to fully replicate the above findings. One reason for this discrepancy might consist of the difference between alcohol drinking and sexual behavior. Issue-involvement in alcohol drinking and in responsible sexual decision-making may actually have a different impact on audiences. People of college age might simply invest more attention into alcohol-related issues, while issues involving STDs might not interest them enough at this stage of their lives.

Whether the lack of support for the second half of Hypothesis 1 lies in the weakness of the stimulus or in the difference between alcohol drinking and sexual health involvement, it shows that future research ought to focus on the counter-argumentative, less involved portion of the audience, in order to improve the creation of effective messages. We know that statistical evidence has an effect, but "preaches to the choir," and does not seem to work with low-involved audience members, though precisely that part of the audience is targeted in health education campaigns. The effectiveness of entertainment-education is undisputed (Singhal & Rogers, 1989; Nariman, 1993) and has emerged from empirical research. Unfortunately, the cognitive information-processing mechanisms behind the effectiveness of entertainment-education messages are yet to be fully comprehended.

The lack of support for Hypothesis 2 probably lies in the weakness of the stimuli. We know from the EE literature (Singhal & Rogers, 1989), that the overt attempt to persuade has a negative impact on the audience. We therefore hoped that the simple introduction of the message as intended either for promotional or for entertainment purposes would cue the audience enough to rate the (identical) messages differently. While this gross, simplistic approach has its value, we probably should have enhanced this "framing" by using stronger stimuli. Maybe, had we claimed that the message was taken from an authoritative newspaper or journal article, a popular website or TV show ("WebMD" versus "Friends," for example), or had been endorsed by a celebrity, the manipulation would have had a stronger impact.

Curiously, H2 was supported when controlling for involvement. When considering involvement, people preferred the anecdotes intended for entertainment, although involvement was not a main effect. This would suggest we take involvement into account when gauging the effects of framing, despite any main effect for involvement, or its interaction with framing or message evaluation. This result might simply be an artifact due to the relatively small sample size: we only considered those audience members who received an anecdote (99 out of a total n = 137), one half of which had received the narrative framed as intended for entertainment, the other half the one framed as intended for promotion. Adding involvement to the equation changed the sample size from 99 to 89.

Hypothesis 3 was not supported, probably for the same reasons the second half of Hypothesis 1 failed. The narrative might just have been too short to allow identification of the reader with the portrayed character. In this case, we cannot expect the gender-homophily between reader and portrayed character to have a significant impact on message evaluation. However, the findings' pointing into the predicted direction is a hopeful sign that greater character development, resulting in easier identification, might yield a gender-homophily effect.

The answer to Research Question 1 showed that people exposed to the statistical message rated the source-credibility higher than those exposed to the narrative. This result is fairly intuitive: the explicitly scientific information provided in the statistical message had to be associated with a knowledgeable and somewhat authoritative source. On the other hand, the narratives might not have been evaluated as based on facts. This is indeed a problem for entertainment-education advocates, who struggle with presenting realistic and plausible, yet fictional messages.

RQ 2, which asked whether the "framing" of the message had an effect on source credibility, did not yield significant results. This, again, might have occurred due to the relatively simple, yet weak manipulation discussed under Hypothesis 2. We would have expected a higher source-credibility for the narrative introduced as intended for entertainment purposes. Again, we might have been able to find an effect by actually introducing the narratives as either taken from WebMD, or from the script of a popular television show (e.g. "Friends").

The answer to Research Question 3 showed that, overall, females rated the messages higher than males. This probably occurred, because young women might be more accustomed to sexual health messages, and they might be socialized to be more responsible in sexual health matters, given the fact that they are the ones who risk unwanted pregnancies. As pharmaceutical contraceptives are prescribed by gynecologists, in occasion of a doctor's visit, women might be more exposed to sexual health messages than men.

A sufficient number of individuals in this study show a high opinion of themselves as decision makers and information users relative to sexual health issues. In some cases these individuals were not able to reveal specific symptoms of fairly common STDs for which they are at high risk. In addition, the majority of respondents who noted HIV/AIDS as the first STD when asked to list some suggests at minimum an agenda-setting media effect, given the media attention to HIV/AIDS, and a possible cultivation effect of a distorted reality where the biggest danger for themselves is acquiring HIV/AIDS and not more realistic health risks, given their demographics.

It is also important to note that the most preferred and trusted sources ranked in this study were not on-line websites, chat rooms and other new information technology sources. This has interesting bearing on where we ought to put our efforts in adolescent health campaigns, including investigations of web credibility concerns.

This study yielded relatively few findings regarding the cognitive processes involved in the processing of entertainment-education messages. It did show that statistical evidence was most appealing to audience members already involved in or agreeing with the content of the message. Yet, far too many health education campaigns are still relying on dry, scientific evidence. This dilemma emphasizes the need for the creation of more effective messages appealing to the counter-argumentative audience members. We have to acknowledge that the narrative used as an instrument in this study did not have the sufficient length for audience members to identify with the character portrayed in the story. Vicarious or parasocial relationships, as described in Bandura's (1977, 1994) social learning theory, cannot be established between the receiver and the character of the story with such a short exposure to the stimulus. However, empirical evidence from entertainment-education research suggests that more engaging types of entertainment-education messages are accepted by broad audiences, and that they even yield measurable behavior changes (decreasing birthrates, increasing condom-sales, and visits to a physician, for example; Nariman, 1993; Singhal & Rogers, 1989). On John Hopkins University's website, a whole archive of completed and ongoing entertainment-education projects shows the ample use of this communication approach across the world. Most projects target audiences in so-called "developing countries," and extend from music videos advocating condom use in Senegal, over a children television show in Ecuador trying to implement environmental responsibility in youngsters, to participatory short film projects in Uganda (http://www.jhuccp.org/ee/archive.stm). The German section of UNICEF's web site features an interactive game, called "Catch the Sperm," which centers around the serious topic of AIDS in a light-hearted way (http://www.unicef.de/catch/index.html). These above examples show how contemporary education efforts increasingly use entertainment in their messages.

Yet, besides of all these anecdotal success reports, still much research has to be conducted to elucidate the cognitive mechanisms of information processing responsible for the success of entertainment-education strategies.

Some researchers (Brown & Singhal, 1994) have reported ethical concerns involved in the field of entertainment-education. Particularly in situations in which the entertainment-education message targets a multicultural, heterogeneous audience, the choice of the advocated behaviors is critical. In multilingual societies like India, even the choice of the language may exclude entire ethnic groups from the benefit of the message. Similarly, sexual issues might encounter different taboo levels in audience members of different religious affiliations. While these concerns have their validity, others, which might compare government-run entertainment-education campaigns to questionable propaganda, do not seem to be justified considering the scope of the problems addressed through entertainment-education strategies. In a globalized world increasingly dominated by corporate brainwash, the attempts to diffuse prosocial behaviors seem morally justified. Historically, the first attempts to employ entertainment-education techniques in campaigns were employed or sponsored by governments of developing countries. Increasingly, also in these countries, commercial television represents a threatening competition to public broadcasting. It is therefore refreshing to note that certain contemporary EE projects have integrated audience's feedback with the actual production of the message. This participatory, grassroots approach not only leads to more effective, tailored messages, but also empowers local populations through this interactive process.

Future research should also concentrate on the effectiveness of single components of entertainment, such as humor, suspense, drama, conflict, etc. Commercial advertising has changed in time, shifting from product-oriented to so-called "lifestyle" ads, which seem to have a more powerful impact on the audience. In the entertainment-education realm, there are some examples of more light-hearted approaches replacing older campaigns focusing on scientific evidence or fear-appeal. Examples are the new "Truth" anti-smoking campaigns on American television, or playful, colorful condom advertisements on billboards in Germany.

References

Archive of Entertainment-Education projects on the website of *John Hopkins Center for Communication Programs*: **http://www.jhuccp.org/ee/archive.stm**

UNICEF's online computer game "Catch the Sperm:" **http://www.unicef.de/catch/index.html**

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1994). Social cognitive theory of mass communication. In J.

Bryant and D. Zillmann (Eds.), *Media effects: Advances in theory and research,* (pp. 61-90), Hillsdale, NJ: Lawrence Erlbaum Associates.

Bentley (1964). *The life of the drama.* New York, NY: Atheneum.

Block, L.G., & Keller, P.A. (1997). Effects of self-efficacy and vividness of health communication. *Journal of Consumer Psychology, 6,* 31-54.

Brosius, H.B., & Bathelt, A. (1994). The utility of exemplars in persuasive communications. *Communication Research, 21,* 48-78.

Brown, W.J. (1992a). Sociocultural influences of prodevelopment soap operas in the Third World. *Journal of Popular Film and Television,* 19*(4)*, 157-64.

Brown, W.J. (1992b). The use of entertainment television programs for promoting prosocial messages. *Howard Journal of Communications,* 3*(3/4)*, 68-80.

Brown, W.J., & Singhal, A. (1994). Ethical considerations of promoting prosocial messages through the popular media. *Journal of Popular Film and Television, 21(3),* 92-99.

Brown, W.J., & Singhal, A. (1999). Entertainment-education media strategies for social change: Promises and problems. In Demers, D., & Viswanath, K. (Eds.), *Mass media, social control, and social change: A macrosocial perspective.* Ames, IA: Iowa State University Press.

Cohen, C.E. (1981). Goals and schemata in person perception: Making sense from the stream of behavior. In N. Cantor & J.R. Kihlstrom (Eds.), *Personality, cognition, and social interaction.* Hillsdale, NJ: Erlbaum.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research.* Reading, MA: Addison-Wensley.

Food and Agricultural Organization (1987). *"The Archers"–an everyday story of country folk.* Rome: Food and Agricultural Organization.

Gerrig, R.J. (1993). *Experiencing narrative worlds: On the psychological activities of reading.* New Haven, CT: Yale University Press.

Green, M.C., & Brock, J.C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Psychology 79,* 701-721.

Higgins, (1987) Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319-340.

Hovland, C.I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly, 106*, 635-650.

Kato, K., & Markus, H. R. (1993). The role of possible selves in memory. *Psychologia, 36*, 73-83.

Libman, E., Rothenberg, I., Fichten, C.S., and Amsel, R. (1985). The SSES-E: a measure of sexual self-efficacy in erectile functioning. *Journal of Sex and Marital Therapy, 11*, 233-247.

McGuire, W.J. (1989). Theoretical foundations of campaigns. In Rice, R.E., & Atkin, C.K. (Eds.), *Public Communication Campaigns* (2nd ed., pp. 43-65). Newbury Park, CA: Sage.

Meyerowoitz, B.E, & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology, 52*, 500-510.

Millstein, S. G., Petersen, A. C., & Nightingale, E. O. (Eds.) (1993). Adolescent health promotion: Rationale, goals, and objectives. In S.G. Millstein, A. C. Petersen, & E. O. Nightingale (Eds.), *Promoting the health of adolescents: New directions for the twenty-first century* (pp.3-12). New York: Oxford University Press.

Nariman, H.N. (1993). *Soap operas for social change: Toward a methodology for entertainment-education television.* Westport, CT: Praeger.

Petty, R.E., & Cacioppo, J.T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* New York, NY: Springer.

Piotrow, P.T. (1994). Entertainment-education: An idea whose time has come. *Population Today,* March, 4-5.

Reeves, B., Chaffee, S., & Tims, A. (1982). Social cognition and mass communication research. In M. E. Roloff, & C. R.Berger (Eds.), *Social cognition and communication* (pp. 287-326). Beverly Hills, CA: Sage.

Reinard, J.C. (1988). The empirical study of the persuasive effects of evidence: The status after fifty years of research. *Human Communication Research, 15*, 3-59.

Rogers, E.M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: Free Press.

Rosenthal, D., Moore, S., & Flynn, I. (1991). Adolescent self-efficacy, self-esteem and sexual risk-taking. *Journal of Community & Applied Social Psychology 1*, 77-88.

Rouner, D., Slater, M.D., Domenech-Rodríguez, M. (2003, in press). Gender-role and sexual imagery in television advertisements: A message and adolescent audience processing analysis comparing beer to non-beer commercials. *Journal of Broadcasting and Electronic Media.*

Seiffge-Krenke, I. (1998). *Adolescents' health: A developmental perspective.* Mahwah, NJ: Lawrence Erlbaum Associates.

Severin, W.J., & Tankard, J.W., Jr. (1997). *Communication theories: Origins, methods, and uses in the mass media* (4th ed.). White Plains, NY: Longman.

Singhal, A., Obregon, R., & Rogers, E.M. (1994). Reconstructing the story of *Simplemente María*, the most popular *telenovela* in Latin America of all time. *Gazette 54*, 1-15.

Singhal, A. & Rogers, E. M. (1989). Prosocial television for development in India. In R.E. Rice, & C. K. Atkin (Eds.), *Public Communication Campaigns* (2nd ed., pp. 331-350). Newbury Park, CA: Sage.

Singhal. A. & Rogers, E. M. (1999). *Entertainment education: A communication strategy for social change.* Mahwah, NJ.: Lawrence Erlbaum Associates.

Slater, M.D. (1997). Persuasion processes across receiver goals and message genres. *Communication Theory, 7*, 125-148.

Slater, M.D., & Rouner, D. (2002). Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory,* 12:2, 173-191.

Slater, M.D., & Rouner, D. (1997). How message evaluation and source attributes may influence credibility assessment and belief change. *Journalism & Mass Communication Quarterly, 73* (4), 974-991.

Slater, M.D., & Rouner, D. (1996). Value-affirmative and value-protective processing of alcohol education messages that include statistical evidence or anecdotes. *Communication Research,* 23 (2), 210-235.

Solomon, W.S. (1989). A social marketing perspective on communication campaigns. In R. E. Rice, & C. K. Atkin (Eds.), *Public Communication Campaigns* (2nd ed., pp. 87-104). Newbury Park, CA: Sage.

Appendix–Three versions of stories (female protagonist anecdote, male protagonist anecdote, statistical story)

For the Rest of Cynthia's Life

When Cynthia entered her junior year of high school, she was diagnosed with an incurable disease, now in epidemic proportions in the United States. She has Herpes Simplex Virus (HSV). Her doctor told her some people don't know they have HSV but she was one of three people of the last 15 patients who have it. During her first outbreak, she went a little late to the doctor and was diagnosed as not having the disease, but her relief was shaken when the lesions recurred and her tests turned positive. She was told blood tests are not definitive and only lesions at an early stage can be accurately diagnosed. Her recurring genital sores are painful and embarrassing, so her doctor advised her to use a horrible medication every day. The most serious problems with taking it are the nausea and the abdominal pains. Within minutes of her swallowing it, her insides scream in pain, although she remains silent on the outside. Sometimes she doubles over and feels like she's going to pass out. She forces herself to take it daily, because it's not effective if she doesn't. It seems to help prevent the symptoms from occurring as frequently, although she still has horrible, unpredictable outbreaks. She had a terrible outbreak of itching, painful blisters the first day of summer camp, the day of her final competition in the state swim meet and when she arrived in New York for her class trip. Cynthia feels like she faces the likelihood of an outbreak during any special event. And coping with the pain and nausea intensifies when she's not in the privacy of her own home. She's never been able to tell her two best friends about all of this. She was not really sexually active before being diagnosed with HSV, but since being diagnosed she is not active at all. She doesn't know whether her old boyfriend is aware she has HSV. She is so afraid she will give the disease to someone, and this can occur even when she's not having an outbreak, according to her doctor. If she ever finds someone she loves, with whom she wants to have sex, she faces telling him about this. And she will likely become infected as well, as one cannot be certain to protect against contracting it. She has completely lost interest in sex or relationships. She also knows that having a child may mean passing this hideous disease onto her children. Her mother's friend told her about a woman with HSV whose child had a birth defect and couldn't hear–the baby may also have been blind. She and all her future family members, if they contract herpes simplex, will be more susceptible to HIV and other sexually transmitted diseases. Not only is she destined to inflict this disease on people she loves, but she has become a victim of daily pain, secrecy and uncertainty, which will last for the rest of her life.

For the Rest of David's Life

When David entered his junior year of high school, he was diagnosed with an incurable disease now in epidemic proportions in the United States. He has Herpes Simplex Virus (HSV). His doctor told him some people don't know they have HSV but he was one of three people of the last 15 patients who have it. During his first outbreak, he went a little late to the doctor and was diagnosed as not having the disease, but his relief was shaken when the lesions recurred and his tests turned positive. He was told blood tests are not definitive and only lesions at an early stage can be accurately diagnosed. His recurring genital sores are painful and embarrassing, so his doctor advised him to use a horrible medication every day. The most serious problems with taking it are the nausea and the abdominal pains. Within minutes of his swallowing it, his insides scream in pain, although he remains silent on the outside. Sometimes he doubles over and feels like he's going to pass out. He forces himself to take it daily, because it's not effective if he doesn't. It seems to help prevent the symptoms from occurring as frequently, although he still has horrible, unpredictable outbreaks. He had a terrible outbreak of itching, painful blisters the first day of summer camp, the day of his final competition in the state swim meet and when he arrived in New York for his class trip. David feels like he faces the likelihood of an outbreak during any special event. And coping with the pain and nausea intensifies when he's not in the privacy of his own home. He's never been able to tell him two best friends about all of this. He was not really sexually active before being diagnosed with HSV, but since being diagnosed he is not active at all. He doesn't know whether his old girlfriend is aware he has HSV. He is so afraid he will give the disease to someone, and this can occur even when he's not having an outbreak, according to his doctor. If he ever finds someone he loves, with whom he wants to have sex, he faces telling her about this. And she will likely become infected as well, as one cannot be certain to protect against contracting it. He has completely lost interest in sex or relationships. He also knows that having a child may mean passing this hideous disease onto his children. His mother's friend told him about a woman with HSV whose child had a birth defect and couldn't hear–the baby may also have been blind. He and all his future family members, if they contract herpes simplex, will be more susceptible to HIV and other sexually transmitted diseases. Not only is he destined to inflict this disease on people he loves, but he has become a victim of daily pain, secrecy and uncertainty, which will last for the rest of his life.

The Other Epidemic: Herpes Simplex Virus

Genital herpes is a disease caused by Herpes Simplex Virus (HSV). It is an epidemic in the United States today. This disease recurs and has no cure. Scientists estimate 45 million people, or 1 out of every 5, in the United States may have it. Some doctors say they detect it in about 3 out of every 15 patients. Some may never recognize they have the infection. Or they may not consider the first outbreak of the disease serious. Studies show 2 out of 3 people with HSV don't know they are infected and contagious. Blood testing doesn't necessarily detect it, and blood tests don't distinguish between types of herpes viruses. So early diagnosis is imperative. If infected people do not get to a doctor in the early stages of an outbreak, cultures of the lesions may show a false negative. Doctors diagnose about ½ of the people with the disease incorrectly during their first outbreak, because they need to do tests before the sores heal. For those who are diagnosed, episodes of the outbreak may occur frequently and unexpectedly. This involves itchy and painful genital sores. For those who have regular outbreaks, a drug can control the symptoms. Some medications that curb outbreak frequency and severity have the unpleasant side effect of induced nausea and sharp abdominal cramps. People with weak immune systems suffer more bouts of the symptoms of HSV. Research on spermicides that are used with condoms and diaphragms shows certain chemicals that might work against the infection from HSV harm the body in serious other ways. One of the most critical problems with the disease is adults inflicting it onto innocent children. About 1 in 3,000 infants nationally get this disease at birth, according to experts at major universities. Further, if a female contracts HSV late in her pregnancy, her baby often suffers from severe brain damage, and sometimes blindness or deafness. Experts believe HSV speeds the spread of HIV, with people infected with HSV experiencing more than 9 times the chance of contracting HIV. Infected people also have a greater likelihood of death if they suffer from lupus or severe burns. Health professionals claim about 35% of their patients have herpes. They recommend abstinence from sex as the only true way to curb the spreading of the disease or risking the chance of getting another debilitating sexually transmitted disease. Many people believe condom use protects from becoming infected with this disease, but there is no safe time for sexual relations without transmitting this disease. Health officials claim even when a person has no visible symptoms, the virus can shed from the skin and infect others. Condoms provide only partial protection because the virus can spread from body parts not covered by condoms. Infected people will suffer from this disease for the rest of their lives unless medical science discovers a way to control it. Scientists hope for a vaccine against the virus in the next decade.

Attributions of Advertising Influence and Negative Stereotypes Among First- and Third-Person Perceptions

By Don Umphrey, Ph.D.
Associate Professor
Southern Methodist University
P.O. Box 750356
Dallas, TX 75275
dumphrey@mail.smu.edu
(214) 768-3370

and

Tom Robinson, Ph.D.
Assistant Professor
Southern Methodist University

Paper Presented to the Commication Theory and Methodology Division, Annual Convention of the AEJMC, Kansas City, 2003.

ABSTRACT

Attributions of Advertising Influence and Negative Stereotypes Among First- and Third-Person Perceptions

Recent high school graduates and university seniors both judged the influence of four magazine advertisements aimed at different age groups on each other, on people in their mid-40s, and on people in their 70s. Both samples demonstrated first-person findings with advertisements for products aimed at younger people and third-person perceptions for people in their mid-40s and 70s with advertisements for products aimed at older individuals. When asked the basis for their judgments of advertising influence of people in their 70s, more than one-third of each sample used negative stereotypes of the elderly.

Attributions of Advertising Influence
and Negative Stereotypes
Among First- and Third-Person Perceptions

Third-person perceptions have been well established in the mass communication literature in the two decades since the introduction of concept by Davison (1983).

The basic idea, based on attribution theory from the field of social psychology (Gunther, 1991; Hoffner et al., 1999; Rucinski & Salmon, 1990), is that when people are exposed to media messages, they make judgments about the effects of that communication on others. When individuals believe that others are more greatly affected than themselves, that is a third-person perception. When individuals believe they are more greatly affected than others, that is a first-person perception (Andsager & White, 2001; Atwood, 1994; Driscoll & Salwen, 1997; Rucinski & Salmon, 1990). A second-person perception (Neuwirth and Frederick, 2002) has been identified to describe situations where in comparison to others, individuals perceive they are equally affected.

When making attributions about the effects of a message on others, individuals may be wrong. After all, without an ability to monitor the cognitive processes of another person, how could anyone really assess the impact of a message on another individual? Incorrect attributions, known as correspondent bias (Jones, 1979) or the fundamental attribution error, generally lean in a certain direction. Namely, attributions of behaviors relating to the self are based on external, situational factors, while attributions of the behaviors of others come from stable personality traits (Jones & Nisbett, 1972; Monson & Snyder, 1977). The fundamental attribution error has been linked to third-person perceptions (Gunther, 1991; McLeod, Detenber & Eveland, 2001; Paul, Salwen & Dupagne, 2000; Rucinski & Salmon, 1990).

Also linked to the fundamental attribution error is stereotyping (Mackie, et al., 1996). This term is defined as "sets of traits ascribed to social groups . . . used to predict and explain behavior" (Stephan, 1985, p. 600). It is well documented that these attributions are often wrong (Bar-Tal, et al., 1989; Lee, Jussim & McCauley, 1995; Macrae, Stangor & Hewstone, 1996). Stereotypes may be positive or negative, but even the positive ones may be harmful by creating an unrealistic, overly optimistic picture (Miller, et al., 1999). Negative stereotypes are associated with prejudice and discrimination (Palmore, 1990).

These two functions of the fundamental attribution error--third-person perceptions and stereotyping--may sometimes merge. In fact, an incident that may serve as an example of such a merger was cited by Davison (1983) as one thing provoking his thinking to formulate the idea of the third-person effect. The incident was an account Davison read about a group of African-American soldiers stationed on the Pacific Island of Iwo Jima during World War II. Learning of their location, the Japanese dropped propaganda leaflets urging these soldiers to either desert or surrender because it was a white-man's war. Despite the fact that this unit had already distinguished itself in battle, it was withdrawn the following day. Apparently, the unit's white officers exhibited third-person perceptions by assuming that the African-Americans might actually attend to the advice of the enemy. Was this assumption about black troops based on negative stereotypes of African-Americans? Probably, but there is no way of actually knowing.

It is the purpose of this paper to further explore the fundamental attribution error at the juncture of third-person perceptions and stereotypes. Namely, when individuals look at an advertisement, what are their perceptions of the effect of this advertising on others, including those who are removed from them by either one or two generations? What attributions are given for such assessments? Will attributions include negative stereotyping according to age? What role does advertising expertise play in perceptions of advertising influence and attributions? To address these questions, this study compared samples of recent high school graduates and university seniors who are advertising majors.

Review of Literature

Advertising Studies

A meta-analysis of third-person research (Paul, Salwen & Dupagne, 2000) revealed relatively few studies focusing on advertising but a majority examining either television or general mass media implications. No study to date has examined third-person perceptions of advertising and stereotyping.

Of those studies dealing with advertising, there have been two major foci.

One of these has been on the biased nature of advertising giving it less credibility and a greater likelihood of producing third-person effects. Such comparisons have been made with editorial matter (Brosius & Engle, 1996; Gunther & Mundy, 1993); public service announcements (Chapin, 1999; Gunther & Thorson, 1992), and news and debates (Rucinski &

Salmon, 1990). Related to these studies on the basis of lower credibility, persuasively weak messages were more apt to yield third-person findings (White, 1997).

The second major focus has been on advertisements with perceived pro- or anti-social content or effects. Topics have included negative political advertising (Cohen & Davis, 1991), tobacco and alcohol products (Banning, 2001), controversial products (Shah, Faber & Youn, 1999), safer sex messages (Chapin, 1999), anti-smoking (Henriksen & Flora, 1999), AIDS prevention (Duck, Hogg & Terry, 1999), body images (David & Johnson, 1998) and teen images (Milkie, 1999).

Not fitting into the above two categories have been two studies on advertising appeals. Third-person effects were associated with neutral but not emotional ads (Gunther & Thorson, 1992), but there was no difference in an emotional ad compared to an informative ad (Duck, Hogg & Terry, 1999).

Stereotyping

In terms of the fundamental attribution error discussed above, the stereotype draws upon perceived stable personality traits that have been attributed to an entire social grouping. Ryan, Park and Judd (1996) concluded that outgroup stereotypes are characterized by greater generalization and exaggeration in comparison to in-group stereotypes. Operario and Fiske (2001) tie stereotyping to "fuzzy sets" of mental prototypes which are "the average of most typical memory of a category" (p. 28), but in most cases, according to the researchers, the prototype does not actually exist.

Pettigrew (1979) proposed "the ultimate attribution error" in which negative acts by out-group members are much more apt to be attributed to dispositional causes than similar acts by in-group members.

Mackie et al., (1996) cited two reasons why individuals stereotype others.

The first is a need for "cognitive efficiency." Trope (1989) called this the " . . . inevitable consequences of the limitations of human information processing capabilities and the complexity of social reality. (p. 134)" Consequently, people are lumped together into social categories on the basis of providing different stimuli than those in other categories, and characteristics of the group are attributed to individuals within that group.

The second reason for stereotyping, according to Mackie et al. (1996), involves self-enhancing judgments based on in-group/out-group categorizations. This has been tied to social identity theory (Tajfel & Turner, 1979) which posits that the self-esteem of individuals is related to maintaining a positive image of the groups to which they belong. Attributions typically lean in the direction of in-group superiority, concluded Hewstone (1990). Ross and Fletcher (1985) wrote that the self-flattering tendencies of individuals is "one of the best established, most often replicated, findings in social psychology" (p. 104).

Parallel third-person findings

This self-serving bias also is documented in the third-person literature (Hoffner, et al., 2001; Lasorsa, 1989; Perloff, 1993) as individuals flatter themselves by believing they are less vulnerable to media impact than others. Lasorsa (1992) noted that this bias may work in this way with low-credibility messages: "I know how biased this message is but others may not recognize its propagandistic nature and therefore, will 'fall for' what I 'see through'" (p. 171-172).

Self-serving biases associated with downward comparisons were used by David and Johnson (1998) to explain findings of higher self-esteem being associated with third-person perceptions.

In-group and out-group membership has also accounted for third-person perceptions (Duck, Hogg & Terry, 1999).

Stereotypes by Age

Among the groups stereotyped are older individuals (Hummert, 1990; Palmore, 1990). Negative stereotypes of older people were found among children ages 10 and 11 (Falchikov, 1990), adolescents (Doka, 1985-86), undergraduates (Levin, 1988; Sanders et al., 1984) and people who were older (Hummert, et al., 1994).

Schmidt & Boland (1986) discovered that elderly people are apt to be perceived as vulnerable and despondent because their lined faces may give them the appearance of being fearful or sad. Even the gait of elderly people leads toward attributions of both weakness and unhappiness (Montepare & Zebrowitz-McArthur, 1988). Pecchhioni and Croghan (2002) described how younger individuals recognize old age cues; these tend to activate negative stereotypes of the elderly and result in notions that communication with such a person would be less satisfying.

In comparison to the way they talked to their peers, care-givers of the elderly were more apt to use more questions and repetitions (Ashburn & Gordon, 1981). Caporael (1981) documented that caregivers of the elderly used "baby-talk" which was very similar in manner to the way they would talk with children. A subsequent study (Caporael, Lucaszewski & Culbertson, 1983) found that caregivers thought that some of the elderly preferred this manner of speech.

Media Influence

In terms of socialization leading to stereotyping, Mackie, et al. (1996) rank the media in western societies as the most powerful influence behind family and friends. These researchers provide evidence showing that beliefs about individuals may be formed on the basis of media stereotypes. Storey (1977) found that adult attitudes toward older people come from depictions of them in literature. In an examination of depictions of the elderly in magazine ads from 1956 to 1996, Miller et al. (1999) found an increasing use of negative stereotypes in more recent years.

Third-person findings and age

No study has previously linked stereotypes of age and third-person findings. However, the issue of age has been addressed. Glynn and Ostman (1988) found a negative relationship between age and perceiving greater media effects on others. Younger people were more apt than older individuals to have third-person perceptions (Driscoll & Salwen, 1997), but the variance was very low. Brosius and Engle (1996) found greater third-person perceptions among people over 40. Neither Rucinski and Salmon (1990) nor Tiedge et al. (1991) discovered a relationship between age and third-person effect.

These ambiguous findings regarding age may indicate that third-person perceptions are not inherent to a particular demographic, per se, but to a combination of a demographic and an issue that is salient to that group. This would be consistent with a speculation of Davison (1983) and subsequent findings (Mutz, 1989; Perloff, 1989; Tiege, et al., 1991; Vallone, et al., 1985). For example, Arabs and Jews produced opposite-direction third-person perceptions of the same news account concerning the Middle East (Perloff, 1989); but it may be speculated that these findings would not be replicated in news about trends in higher education or any number of other issues.

Social distance

Perceptions of social distance may play a part in third-person perceptions according to age. Gunther (1995) concluded that the third-person effect was more apt to occur when others were perceived as different from themselves.

Social distance was related to third-person findings in the Duck, Hogg and Terry (1999) study and was explained in terms of in-group/out-group membership.

Research using social distance to explain third-person perceptions have included people who were not on campus in comparison to those who were (Mutz, 1989; Wu & Koo, 2001), others in the U.S. compared to those in a local community (Hoffner et al., 2001), "other peers" compared to "best friends" of children (Henriksen & Flora, 1999), and "others" compared to "best friends" (Chapin, 2000). Cohen et al. (1988) discovered linear increases in third-person perceptions as social distance increased from other students on the same campus to other people in the same state to the general public at large." White (1997) reported similar linear findings.

Education has been tied to the idea of social distance and resulted in increased third-person effects (Peiser & Peter, 2000; Tiedge, et al., 1991).

If education is related to social distance, what about knowledge or perceived knowledge of issues? There have been mixed findings concerning third-person perceptions dealing with either knowledge or perceived knowledge (Atwood, 1994; Driscoll & Salwen, 1997; Lasorsa, 1989; Price, Huang & Tewksbury, 1997; Price & Tewksbury, 1996), leading Price and Tewksbury to conclude that personal importance of an issue may outweigh either actual or perceived knowledge.

Naive theories of media effects

The roots of the third-person tradition in attribution theory are reflected in a variable that intercedes to override third-person perceptions based on social distance. Before proceeding with this line of reasoning, it may be instructive to examine some basic concerns of attribution theory.

Attribution theory was introduced by Heider (1958). His perspective is known as "naive psychology" (Folkes, 1988; Kelley & Michela, 1980). Focusing on interpersonal relationships, Heider believed people were "naive psychologists" who sought common-sense answers to understand the world around them and the behavior of others (Mizerski, Golden & Kernan, 1979; Weiner, 1990). Attribution theory is "based on the conviction that if we can capture the naive understandings of the person on the street, we can accurately infer . . . his other expectations and actions" (Jones, 1985, p. 89). Tying this idea to third-person findings, Price, Huang and Tewksbury (1997) observed, " . . . people possess common-sense theories about media" that "organize expectations about and reactions to specific media messages" (p. 527).

In a study of college students and the perceived effects of rap and death metal music, Eveland, et al. (1999) speculated that a "naive theory of media effects" exerted a stronger influence on third-person findings than social distance. The students demonstrated third-person effects for people who were their own age but not for people who were ages 40 and older. The students (probably correctly) believed that the older group was unlikely to be exposed to the music in question, according to the authors. When it comes to estimating media effects, according to these researchers, people may draw upon "the most naive theory. . ." This theory and its elements were identified by McLeod, Detenber and Eveland (2001) as a version of the "magic bullet" theory with the idea that greater effects go hand-in-hand with greater exposure.

The two studies cited below have demonstrated that naive theories of media effects intervened to alter findings related to third-person perceptions. Each is similar to Eveland et al, (1999) in dealing with perceived age-appropriateness of specified products.

Third-person perceptions were associated with a desire to censor pornography to adults (Wu & Koo, 2001), but the third-person perception applied to children was not associated with a desire to censor. This was apparently due to a belief that children would not have access to it.

There were significant relationships between third-person perceptions and a desire to censor handgun, gambling and beer web sites when adults perceived other adults (Youn, Wan & Faber, 2001), but there was a non-significant relationship on the same measure when adults perceived teens. The authors speculated that the difference may be due to the fact that adults believed teens would have fewer opportunities to purchase those products.

Hypothesis and Research Questions

H1 The naive theory of media effects will be reflected in perceptions about target audiences (or age appropriateness of certain products). As naive theories of media effects exert a stronger influence than perceptions of social distance, perceptions of the effects of advertising on others will vary according to the product being advertised.

Common sense could lead to speculation that in comparison to recent high school graduates, university seniors who are advertising majors would exhibit greater understandings about the target audiences of products. However, a clear picture does not emerge when examining the literature in regards to a combination of the following: a) education as a measure of social distance; b) the effects of actual and/or perceived knowledge, which have never been conceptualized in terms of a major area of study at a university, and c) the influence of issue salience, which could be interpreted as advertising issues being salient to advertising majors and/or the perceptions of potential audiences of advertising as they are exposed to an advertised product or advertising material that is salient to them.

RQ1 What differences, if any, will emerge when comparing perceptions of advertising influence between recent high school graduates and senior advertising majors?

There is only one study in the third-person tradition to serve as a guide as to the attributions of media influence: McLeod, et al. (2001) speculated that the "naive theory" was the notion that greater effects go hand-in-hand with greater exposure. Seeking to build on that finding, this study proposes the following research questions:

RQ2a After noting their perceptions of media influence on others, what attributions will be given by respondents for their perceptions?

RQ2b Will ratios of attributions by category vary between recent high school graduates and senior advertising majors?

RQ3a To what extent will attributions of the elderly include negative stereotypes, if any?

RQ3b If there are negative stereotypes of the elderly, will they differ between recent high school graduates and university seniors who are advertising majors?

Methodology

Two samples were given self-administered questionnaires that had been pre-tested. One was a convenience sample of senior advertising majors from 13 universities located in the southwest and mid-west U.S. during the spring and summer months of 2002. The other consisted of recent high school graduates who were in attendance at a freshman orientation program at a private university in the southwest in mid-August, 2002; these respondents were in one of two auditoriums on the university campus to which they were randomly assigned to wait for transportation to a retreat.

The average time to complete the questionnaire was 4 minutes, 48 seconds for recent high school graduates and 4 minutes, 41 seconds for senior advertising majors.

On the cover of the questionnaire, respondents in each sample were asked their gender and age. Additionally, the cover sheet for the university advertising majors also asked their major and year in school. Questionnaires were discarded when the survey instrument was filled in by someone other than a senior advertising major.

Instructions on the cover sheet read as follows: "We are interested in the opinions of students about aspects of advertising. Your responses are confidential and will not/cannot be traced back to you." (Students were asked to not put their names or any other identifying marks on the instrument.)

Following the demographic questions, the instrument read,

> In the pages below, you will find four advertisements. Following each advertisement is a page asking you to judge the influence of that advertisement on yourself and on other people of different ages. After you have completed the evaluations of the advertisements, please complete the questions on the last page.

The ads were each four-color, full-page advertisements that had appeared in a magazine, and the ads featured products that would appeal to different age groups. To prevent ordering effects, the advertisements were collated in every possible combination an equal number of times.

The advertised product and a brief description of each of the four ads follow:

Depend Fitted Briefs. A headline at the top of the page reads, "Introducing the driest Depend Briefs ever. Because there's nothing like spending an entire afternoon with old friends." The artwork which dominates the page features two smiling women who could be in their sixties and appear to be having a good time while talking.

Lincoln LS. Artwork on the top half of the ad is a photo of skyscrapers in New York city. The type on this artwork reads. "Shopping on Madison Avenue. Tickets to a Broadway show. And a hot dog in the park. Define luxury for yourself." Artwork on the bottom half of the ad depicts a black Lincoln being driven in a large city. It contains small type describing attributes of the Lincoln LS.

Propel. The full-page artwork depicts a bottle of the product with a bolt of lightening entering the top of it. Superimposed over the bottle is the headline, "It's how Gatorade does water."

Remdex. The artwork features something resembling a yellow traffic sign with a nose depicted on it. Beneath each of the nostrils is an arrow pointing downward, away from the nose. The headline reads, "Take Remdex at the first sign of a cold."

Following each of the four advertisements was a full page asking students about the "influence" of the foregoing ad. "Thinking about the foregoing _________ advertisement . . ." "Overall, how much would you say that YOU are influenced by this advertisement." Below was a scale ranging from 0 (with words beneath it reading "Not at all") to 10 "a great deal;" this scale was used by Eveland, et al. (1999). After this, the same wording and scale were used to ask about "Recent High School Graduates About to Enter College" or "University Seniors Who Are Majoring in Advertising," depending on the sample. Then both groups were asked about "People in Their Mid-40s" and "People in their 70s," again with the same scale.

Instructions on the last page of the questionnaire stated,

> You have just made evaluations as to how advertising might influence different groups of people. What ideas/experience/information, etc., did up draw upon to estimate the influence of the advertising on:

"Recent high school graduates about to enter college?" or "University seniors who are majoring in advertising?"

"People in their mid-40s?"

"People in their 70s?"

The three open-ended items were coded into three (non-exclusive) categories on a nominal basis as to whether the attribution fit into the category. Categories included: a) respondents mentioned attributes of people in the specific age group; b) respondents made mention of specific people in the age group; c) respondents said things about the product, advertisement, or mentioned the product meeting the needs of a specific audience. Inter-coder reliability (Hosti, 1969) equals 95.7%.

In a few instances where respondents put an answer that did not fit the above three categories, it was because the response was not appropriate to the question. For example, one respondent wrote either "yes" or "no" for each of the open-ended questions. These were counted as missing variables. There were many more cases where respondents left blank one or more of the open-ended questions; these, too, were counted as missing variables.

Usable responses to the open-ended items were as follows: the recent high school graduates, 85.7% for ad majors, 83.3% for people in their mid-40s, 83.9% for people in their 70s; the senior advertising majors, 84.3% for recent high school, 81.9% for people in their mid-40s, 83.1% for people in their 70s.

The open-ended responses pertaining to people in their 70s were analyzed on a nominal basis as to whether they included negative stereotypes of the elderly. Criteria for these negative stereotypes came from both Miller et al. (1999) and Robinson et al. (2002). Intercoder reliability equals 93.5%.

Statistical analyses were conducted using SPSS 10.0 for Macintosh.

Results

Responding to questionnaires were 168 recent high school graduates and 166 senior advertising majors. The sample of recent high school graduates was 46.5% male, 53.0% female, and .6% (1 respondent) who did not indicate gender; the ages ranged from 17 to 20 with a mean age of 18.07 years. The mean age of senior advertising majors at the university level was 22.64 years with a range of 20 to 37; this sample included 31.3% males; 65.7% females, and 3% who did indicate gender.

H1

As can be seen on Table 1, with the Depend ad the recent high school graduates demonstrated third-person perceptions that were linear with age. These findings could be interpreted as respondents recognizing the product is aimed at the elderly, but they could result from social distance.

The remaining findings among the recent high school graduates show a recognition of target audiences over-riding the notion social distance. For example, these respondents rated the least impact on themselves and the greatest on people in their mid-40s with the Lincoln ad. There was a first-person perception with the Propel ad with the elderly estimated as least influenced, while with the Remdex ad the greatest estimated influence was quite close between individuals in their mid-40s and mid-70s.

As demonstrated on Table 2, the senior ad majors had results similar to the recent high school graduates. There were linear increases by age for the influence of Depend but linear decreases by age for Propel. The greatest influence for the Lincoln ad was attributed to individuals in their mid-40s, while scores for Remdex were nearly identical for mid-40s and mid-70s.

H1 is supported. This is the idea that naive theories of media effects (in terms of recognition of age appropriateness of products) override the notion of social distance.

RQ1

When comparisons were made with assessments of advertising influence between the two samples, seven of the 16 measures yielded significant t-tests. (See Table 3.) All but one of these were differences in the way recent high school graduates viewed the senior ad majors or vice-versa. For example, the recent high school graduates rated the effect of the Depend ad greater on themselves than the way the senior ad majors rated them. The recent high school students also rated the effect of the Depend ad on the senior ad majors significantly higher than the ad majors perceived its effect on themselves.

The one significant difference pertaining to the older age groups was with the Lincoln ad on people in their 70s. Thus, the recent high school graduates were in accord with the senior advertising majors on seven of the eight measures dealing with people in their mid-40s and 70s.

In 15 of the 16 measures, the senior ad majors had smaller standard deviations. These tighter clusters around the mean are probably due to the effects of advertising education.

RQ2a and 2b

Examples of respondents mentioning aspects of people in a specific age group:

"These people are more interested in buying things for college--that's the only thing on their mind. They want party, fun, and something to make them think of college." (an ad major assessing a recent high school graduate)

"High school grads are all about themselves and what they think is cool." (an ad major regarding recent high school graduate)

"Want luxury, recognition." (An ad major on people in their mid-40s)

"Older people, especially on tv." (a recent high school graduate on people in their 70s)

Examples of respondents mentioning a specific person in that age group:

"I was there. Been there, done that." (an ad major on recent high school students)

"I thought how my parents would respond." (a recent high school graduate on people in mid-40s)

"The likelihood of my grandparents to use the product." (an ad major on people in their 70s)

Examples of respondents mentioning the product, advertisement, or target audience:

In referring to any age group, many from both samples mentioned target audience concerns, such as this response from an ad major, *"Stuff that fits their specific needs."*

High school grad, *"lots of graphics;"* mid-40s, *"more serious ads;"* people in 70s, *"the people in the ad."* (an ad major)

"The pictures of the old women in the Depend ad definitely attract those in their 70s." (an ad major)

Following are examples of respondents giving answers that were coded in both the category of aspects of people and the category for the ad, product, target:

"This group does not spend near the amount of money other groups do. However, products such as Depend really target toward this group." (an ad major on people in their 70s)

"People in their 70s like to drive huge cars, and they probably need diapers more than another market. Also, they like to buy American (Pearl Harbor, you know). So they'd be influenced by an ad for a big American car." (an ad major on people in their 70s)

Frequencies for the three non-mutually exclusive categories of attributions of the recent high school graduates are found on Table 4, while the same measure for senior ad majors is on Table 5.

In their attributions for each age grouping, around half of the recent high school graduates who responded made attributions relating to aspects of people. Similarly, mentions of the product, ad or targeting also were made by around half of those who made attributions for each of the different age groups. A much smaller ratio made references to specific people. When the recent high school graduates made attributions of senior advertising majors, all but four of those who mentioned attributes of the people in this group (equaling 50% of those who made any attribution) mentioned expertise in advertising. (e.g. *"Are probably very critical of all ads and see the details others may not see."* *"They know all the tricks of advertising so they can see threw* (sic) *these ads."* *"They are looking at the design asspect* (sic) *and how it will influence decisions."*)

When estimating advertising influence on recent high school graduates, 64.3% of the senior advertising majors mentioned aspects of the product, ad or targeting. Other than that, the ratios coded in category were about the same as the recent high school graduates.

Since respondents in the two groups were not asked to make attributions about themselves, age groupings of attributions that were common to the two samples were people in their mid-40s and people in their 70s. Chi square analyses revealed no significant differences between the samples and the ratio of attributions in any category.

Therefore, just as there were minimal differences between recent high school graduates and senior ad majors in third-person perceptions of people in their mid-40s and 70s, there was not much difference in their ratios of attributions, at least as it is revealed in the comparisons in the three coded categories. Of course, there were qualitative differences in their attributions, as demonstrated by the rate of recent high school graduates citing advertising expertise of the senior advertising majors.

RQ3a and 3b

Stereotypes were found among the attributions of advertising influence in the category of characteristics of people.

Examples of attributions of elderly without negative stereotypes (two of which may contain positive stereotypes):

"Older people pay more attention." (recent high school graduate citing a positive stereotype)

(Older people like) *"pretty things."* (recent high school graduate)

"Observation of retired." (senior ad major)

Examples of attributions of elderly with negative stereotypes:

"Don't really care. They will buy most stuff if it has a good ad." (recent high school graduate)

"They are like little kids." (recent high school graduate)

"Are older and more prone to believe anything they hear." (recent high school graduate)

"Usually older people are set in their ways and are not interested in new products." (senior ad major)

"Stereotypical views of the elderly. Image that came to mind was Grandpa Simpson." (senior ad major)

Even though the focus of this paper concerns negative stereotypes of the elderly, it is worth mentioning that the use of stereotypes was not confined to this age group. For example, a senior ad major described recent high school graduates as having short attention spans. There were also numerous negative stereotypes of people in their mid-40s.

Out of the 279 respondents from both samples coded for making attributions, 35.5% (99) included negative stereotypes of elderly people. By sample, 31.2% of the recent high school graduates made attributions of negative stereotypes as did 39.9% of the senior ad majors.

Analysis between the samples was not significant at .05 ($X^2 = .2.3$, d.f. = 1, $p < .083$). An analysis by age was marginally significant (stereotype mean = 20.8, non-stereotype mean = 20.1, $p < .053$).

In an analysis of combined samples, only 2 of the 99 respondents who used negative stereotypes also used an attribution about a specific person. This converts to 8.7% (2 of 25) of those naming a specific person using negative stereotypes ($X^2 = 9.1$, d.f. = 1, $p < .002$).

Discussion

This study demonstrates that in some instances, third-person perceptions are tied to attributions of media influence including negative stereotypes.

Previous research associated third-person perceptions with various conceptualizations of social distance. An exception was discovered by Eveland et al. (1999) when college students did not believe that people in their 40s would be exposed to rap and death metal music. This naive theory of media exposure negated third-person findings based on social distance. Using advertising aimed at various age groups, this study also found that naive theories of media effects based on perceptions of age-appropriateness of certain products exerted a stronger influence than social distance in producing third-person findings.

McLeod, Detenber and Eveland (2001) identified a naive theory of media effects as "the magic bullet theory," an idea that greater influence goes hand-in-hand with greater levels of exposure. Findings here are in accord with that earlier study but add details about the cognitive processes involved. The naive theory varies among individuals. Some make attributions based on characteristics of people of different ages which are stored in their memories. Among these characteristics were more than a third of both samples citing negative stereotypes in regards to people in their 70s. Others cited specific people they knew who fit into a group. Aspects of the product, advertising or the concept of target audience were named by yet other respondents.

When estimating advertising influence on those who were in their mid-40s and 70s, there were minimal differences between recent high school graduates and senior advertising majors. While it might be expected that senior advertising majors would be more apt to make attributions to aspects of the product, ad or target audience, there were no measured differences between the two samples in the ratios of attributions by category. Neither was there a difference in the ratio who used negative stereotypes for attributions. These findings could be interpreted that actual knowledge of advertising did not account for differences in third-person perceptions. But based on earlier cited literature, it is more likely that the mental mechanisms that lead to recognizing target audiences and making attributions that sometimes include negative stereotypes are already in place by the time someone graduates from high school. However, the effects of advertising education were probably manifested in greater unanimity of estimates that were demonstrated by the standard deviations of the advertising majors.

The finding of significantly less stereotyping among individuals who named a specific person in his or her 70s (such as a grandparent) is consistent with previous research. Brewer (1996) concluded that individuating information such as a social relationship reduces category-based impressions. This individuating information could be an exemplar, someone in the category with whom they are/were personally acquainted (Kahneman & Miller, 1986; Stapel & Koomen, 1998) or in using a self-schema (Markus, Smith & Moreland, 1985) (e.g. "I remember myself at that age.") In this study respondents sometimes cited an exemplar when making attributions about people in their 70s, and some of the advertising majors used a self-schema in their attributions of recent high school graduates. Related to the foregoing, another study (Pecchioni & Croghan, 2002) found that closeness of a grandparent served as a mediating factor in stereotyping of older adults.

The information in the above paragraph may dovetail with Tewksbury (2002) conclusions that third-person findings based on social distance increase as the size of the group increases (e.g. from other students at this university to other people in this state). By drawing upon specific individuals within the larger group such as with the exemplar or self-schema, it may be speculated that individuals cognitively decrease group size.

A methodological concern expressed by Neuwirth and Frederick (2002) relating to the use of general estimates of media influence (in comparison to more specific measures of influence) found verification in this study. Here, a general measure was used as respondents were asked about the "influence" of the four advertisements. Concerning such a general measure, Neuwirth and Frederick observed, "we as researchers typically cannot say with any great certainty what kind of media influence is considered by the respondent (p. 115)."

This study demonstrated that respondents from both samples had fairly uniform ideas about the age appropriateness of certain products. With this in mind, it is highly probable that both the senior advertising majors and recent high school graduates perceived "influence" of the Depend ad on people in their seventies in terms of the product being age appropriate. However, they were probably thinking of "influence" in a different way when judging the effects of the same ad on people who were in either their late teens or early 20s. Testimony to this comes from the fact that 50% of the recent high school graduates who made any attribution of senior ad majors cited advertising expertise in some form; thus, the younger students may have judged "influence" in terms of whether ad majors would think the advertising execution was good or bad.

Confusion concerning the meaning of "influence" may help to explain why all but one of the differences between the samples in their perceptions of other age groups were on their perceptions of each other. Certainly, this consideration about the meaning of the word "influence" confounds a notion that differences between the recent high school graduates and the advertising majors was based on social distance created by advertising education.

Another methodological concern is that the Depend ad may have served as a cue to more negative images of the elderly. (e.g. One senior advertising major cited "incontinence" as a characteristic of people in their 70s.) Even though the four ads used in this study were placed in different positions in the survey packet, the page asking for attributions was, by necessity, always last. To test for such cueing effects, future efforts could include some survey packets with only positive images of a stereotyped group and others with both positive and negative.

An oft-noted shortcoming of studies employing surveys is that they are reliant on the vagaries of respondent self-reports. This applies to an even greater extent in this study because the open-ended questions may have prompted respondents to delve into previously unexplored cognitive regions, namely the hows and whys for making attributions.

In comparison to the methodology used here, in-depth interviews would undoubtedly yield a much broader picture of the cognitive processes involved in making attributions of media influence and stereotyping. And rather than relying on self-reports of any kind, would it be possible to use behavioral measures?

Paul, Salwen and Dupagne (2000) noted that skewed results may have been produced by the frequent use of student samples in studies dealing with third-person effects. While the convenience sample of senior advertising majors in this study allowed scrutiny of a group containing special information, it would be useful to test these propositions in a random sample of the public at large.

Perhaps the greatest contribution of this study is that it ties together third-person findings and stereotyping. It is meaningful that here, the negative stereotyping results came from open-ended questions that did not contain any prompts concerning stereotyping. It would be helpful for future studies to use different types of measures of stereotyping.

While this study focused on stereotypes of the elderly, the topic of third-person (or first- or second-person) effects related to stereotyping (and related attitudes and behaviors that may result from it) suggests many possibilities for future research.

References

Andsager, J. L. & White, H. A. (2001). Message credibility and congruence in first-and third-person estimations. Paper presented at the annual convention of the Association for Education in Journalism and Mass Communication, Washington, D. C.

Ashburn, G. & Gordon, A. (1981). Features of a simplified register in speech to elderly conversationalists. International Journal of Psycholinguistics, 8, 7 -31.

Atwood, L. E. (1994). Illusions of media power: The third-person effect. Journalism Quarterly, 71, 269-281.

Banning, S. A.(2001). Do you see what I see?: Third-person effects on public communication through self-esteem, social stigma, and product use. Mass Communication and Society, 4, 127-147.

Bar-Tal, D.. Graumann, C. F., Kruglanski, A. W. and Stroebe.W. (Eds.), (1989). Stereotyping and prejudice: Changing concepts. New York: Springer-Verlag.

Brewer, M. B. (1996). When stereotypes lead to stereotyping: The use of stereotypes in person perception. In In C. N. Macrae, C. Stangor & M. Hewstone (Eds.) Stereotypes and stereotyping (pp. 254-275). New York & London: The Guilford Press.

Brosius, H. D. & Engel, D. (1996). The causes of third-person effects: Unrealistic optimism, impersonal impact, or generalized negative attitudes toward media influence? International Journal of Public Opinion Research, 8, 142-162.

Caporael, L. (1981). The paralanguage of caregiving: Baby talk to the institutionalized aged. Journal of Personality and Social Psychology, 40, 876-884.

Caporael, L., Lucasziewski, M P. & Culbertson, G.H. (1983). Secondary babytalk: Judgments by institutionized elderly and their caregivers. Journal of Personality and Social Psychology, 44, 746-754.

Chapin, J. R. (1999). Third-person perception and sexual risk taking among minority 'at risk' youth. Mass Communication and Society, 2, 16-173.

Cohen, J. & Davis, R. G. (1991). Third-person effects and the differential impact in negative political advertising. Journalism Quarterly, 68, 680-688.

David, P. & Johnson, M. A. (1998). The role of self in third-person effects about body image. Journal of Communication, 48, 37-58.

Davison, W. P. (1983). The third-person effect in communication. Public Opinion Quarterly, 47, 1-15.

Doka, K. J. (1985-86). Adolescent attitudes and beliefs toward aging and the elderly. International Journal of Aging and Human Development, 22, 173-187.

Driscoll, P. D. & Salwen, M. B. (1997). Self-perceived knowledge of the O. J. Simpson trial: Third-person perception and perceptions of guilt. Journalism & Mass Communication Quarterly, 74, 541-556.

Duck, J. M., Hogg, M. A. & Terry, D. J. (1999). Social identity and perceptions of media persuasion: Are we always less influenced than others? Journal of Applied Social Psychology, 29, 1879-1899.

Eveland, W. P., Jr., Nathanson, A. I., Detenber, B H. & McLeod, D. M. (1999). Rethinking the social distance corollary: Perceived likelihood of exposure and the third-person perception. Communication Research, 26, 275-302.

Falchikov, N. (1990). Youthful ideas about old age: An analysis of children's drawings. International Journal of Aging and Human Development, 31, 79-99.

Folkes, V. S. (1988). Recent attribution research in consumer behavior: A review and new directions. Journal of Consumer Research, 14, 548-565.

Freedman, J. L. (1964). Involvement, discrepancy, and change. Journal of Abnormal and Social Psychology, 69, 290-295.

Gilovich, T., Jennings, D. & Jennings, S. (1983). Causal focus and estimates of consensus: An examination of the false consensus effect. Journal of Personality and Social Psychology, 45, 550-559.

Glynn, C. J. & Ostman, R. E. (1988). Public opinion about public opinion. Journalism Quarterly, 65. 299-306.

Gunther, A. (1991). What we think others think: Cause and consequence in the third-person effect. Communication Research, 18, 355-372.

Gunther, A. C. & Mundy, P. (1993). Biased optimism and the third-person effect. Journalism Quarterly, 70, 58-67.

Gunther, A. C. & Thorson, A. (1992). Perceived persuasive effects of product commercials and public service announcements. Communication Research, 19, 574-596.

Heider, F. (1958). The Psychology of Interpersonal Relations. New York: John Wiley & Sons.

Henriksen, L. & Flora, J. A. (1999). Third person perception and children: Perceived impact of pro- and anti-smoking ads. Communication Research, 26, 643-665.

Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup casual attribution. European Journal of Social Psychology, 20, 311-335.

Hoffner, C., Buchanan, M., Anderson, J. D., Hubbs, L. A., Kamigaki, S. K., Kowalczyk, L., Pastorek, A., Plotkin, R .S. & Silberg, K. J. (1999). Support for censorship of television violence: The role of the third-person effect and news exposure. Communication Research, 26, 726-742.

Hoffner, C., Plotkin, R. S., Buchanan, M., Anderson, J. D., Kamigaki, S. K., Hubbs, L. A., Kowalczyk, L., Silberg, K. & Pastorek, A. (2001). The third-person effect in perceptions of television violence. Journal of Communication, 51, 283-299.

Hosti, O. (1969). Content analysis for the social sciences and humanities. Reading, MA: Addison-Wesley.

Hummert, M. L. (1990). Multiple stereotypes of elderly and young adults: A comparison of structure and evaluations. Psychology and Aging, 5, 182-193.

Hummert, M. L., Garstka, T. A., Shaner, J. L., & Strahm, S. (1994). Stereotypes of the elderly held by young, middle-aged and elderly adults. Journal of Gerontology: Psychological Sciences, 49, P240-P249.

Jones, E. E. (1985). Major developments in social psychology during the past five decades. In G. Lindzey & E. Aronson (Eds.) Handbook of Social Psychology, (3rd ed., Vol. 1, pp. 47-107). New York: Random House.

Jones, E. E. (1979). The rocky road from acts to dispositions. American Psychologist, 34, 104-117.

Jones, E. E. & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.) Attribution: Perceiving the Causes of Behavior. (pp. 79-94). Morristown, NJ: General Learning Press.

Kahneman, D. & Miller D. T. (1986). Norm theory: Comparing reality to its alternatives. Psychological Review, 93, 136-153.

Kelley, H.H. & Michela, J. L. (1980). Attribution theory and research. In M. R. Rosenzweig and L. W. Porter (Eds.) Annual Review of Psychology. Vol. 31 (pp. 457-501). Palo Alto: Annual Reviews, Inc.

Lasorsa, D.L. (1989). Real and perceived effects of "Amerika." Journalism Quarterly, 66, 373-378, 529.

Lasorsa, D. L. (1992). How media affect policy-makers: 'The third-person effect.' In J. D. Kennamer, (Ed.), Public opinion, the press and public policy (pp. 163-175). New York: Praeger.

Lee, Y-T, Jussim, L. J. & McCauley, C. R. (Eds.) (1995). Stereotyping accuracy: Toward appreciating group differences. Washington, D.C.: American Psychological Association.

Levin, W. C. (1988). Age stereotyping: College student evaluations. Research on Aging, 10, 138-148.

Mackie, D. M., Hamilton, D. L., Russkind, J. & Rosselli, F. (1996). Social psychological foundations of stereotype formation. In C. N. Macrae, C. Stangor & M. Hewstone (Eds.), Stereotypes and stereotyping (pp. 41-78). New York & London: The Guilford Press.

Macrae, C. N. Stangor, C. & Hewstone M. (Eds.) 1996. Stereotypes and stereotyping (pp. 41-78). New York & London: The Guilford Press.

Markus, H., Smith, J. Moreland, R. L. (1985). Role of the self-concept in the perception of others. Journal of Personality and Social Psychology, 35, 63-78.

McLeod, D. M., Detenber, B. H. & Eveland, Jr., W. P. (2001). Behind the third-person effect: Differentiating perceptual processes for self and others. Journal of Communication, 51, 678-695.

Milkie, M. M. (1999). Social comparisons, reflected appraisals, and mass media: The impact of pervasive beauty images on black and white girls' self-concepts. Social Psychology Quarterly, 62, 190-210

Miller, P. N., Miller, D. W., McKibbin, E. M. & Pettys, G. L. (1999). Stereotypes of the elderly in magazine advertisements 1956-1996. International Journal of Aging and Human Development, 49, 319-337.

Mizerski, R. W., Golden, L. L. & Kernan, J. B. (1979). The attribution process in consumer decision making. Journal of Consumer Research, 6, 123-140.

Monson, T. C. & Snyder, M. (1977). Actors, observers, and the attribution process: Toward a reconceptualization. Journal of Experimental and Social Psychology, 13, 89-111.

Montepare J.M. & Zebrowitz-McArthur, L. (1988). Impressions of people created by age-related qualities of their gaits. Journal of Personality and Social Psychology, 55, 547-556.

Neuwirth, K. & Frederick, E. (2002). Extending the framework of third-, first and second-person effects. Mass Communication & Society, 5, 113-140.

Operario, D. & Fiske, S. T. (2001). Stereotypes: Content, structures, processes and context. In R. Brown & S. L. Gaertner (Eds.) Blackwell Handbook of Social Psychology: Intergroup Processes,(pp. 22-44). Malden, MA: Blackwell Publishers.

Palmore, E. B. (1990). Ageism: Negative and Positive. New York: Springer.

Paul, B., Salwen, M .B. & Dupagne, M. (2000). The third-person effect: A meta-analysis of the perceptual hypothesis. Mass Communication & Society, 3, 57-85.

Pecchioni, L. L. & Crogham, J. M. (2002). Young adults' stereotypes of older adults with their grandparents as the targets. Journal of Communication, 52, 715-730.

Peiser, W. & Peter, J. (2000). Third-person perception of television-viewing behavior. Journal of Communication, 50, 25-45.

Perloff, R. M. (1989). Ego-involvement and the third person effect of televised news coverage. Communication Research, 16, 236-262.

Perloff, R. M. (1993). Third-person effect research 1983-92: A review and synthesis. International Journal of Public Opinion Research, 5, 167-184.

Pettigrew, T.W. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. Personality and Social Psychology Bulletin, 5, 461-476.

Price, V., Huang, L. N. & Tewksbury, D. (1997). Third-person effects of news coverage: Orientations toward media. Journalism & Mass Communication Quarterly, 74, 525-540.

Price, V. & Tewksbury, D. (1996). Measuring the third-person effect of news: The impact of question order, contrast and knowledge. International Journal of Public Opinion Research, 8, 120-141.

Robinson, T., Popovich, T., Gufstafson, B. & Fraser, C., (2002). Seniors' perceptions of seniors stereotypes featured in magazine advertisements: A Q method analysis. Proceedings of the 2002 Conference of the American Academy of Advertising. Jacksonville, FL, p 117-127.

Ross, M. & Fletcher, G. J. O. (1985). Attribution and social perception. In G. Lindsey and E. Aronson (Eds.), Handbook of Social Psychology, 3rd Ed. (pp. 73-122). New York: Random House.

Rucinski, D. & Salmon, C. T. (1990). The 'other' as the vulnerable voter: A study of the third-person effect in the 1988 U.S. presidential campaign. International Journal of Public Opinion Research, 2, 345-368.

Ryan, C.S., Park, B. & Judd, C. M. (1996). Assessing stereotype accuracy: Implications for understanding the stereotyping process. In C. N. Macrae, C. Stangor & M. Hewstone (Eds.), Stereotypes and stereotyping (pp. 121-157). New York & London: The Guilford Press.

Sanders, G. F., Montgomery, J. E., Pittman, J. F., & Balkwell, C. (1984). Youth's attitudes toward the elderly. Journal of Applied Gerontology, 3, 59-70.

Schmidt, D. R. & Boland, S. M. (1986). Structure of perceptions of older adults: Evidence for multiple stereotypes. Psychology and Aging, 1, 255-260.

Shah, D. V., Faber, R. J. & Youn, S. (1999). Susceptibility and severity: Perceptual dimensions underlying the third-person effect. Communication Research, 26, 240-267.

Stapel, D. A. & Koomen, W. (1998). When stereotype activation results in (counter) stereotypic judgments: Priming stereotype-relevant traits and exemplars. Journal of Personality and Social Psychology, 34, 136-163.

Stephan, W. A. (1985). Intergroup relations. In G. Lindzey & E. Aronson (Eds.) The handbook of social psychology, 3rd edition. Reading, MA: Addison-Wesley.

Storey, D. C. (1977). Gray power: An endangered species. Social education, 41, 528-533.

Tajfel, H. & Turner, J.C. (1979). An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (Eds.), The social psychology of intergroup relations (pp. 33-47). Monterey, CA: Brookes-Cole.

Tewksbury, D. (2002). The role of comparison group size in the third-person effect. International Journal of Public Opinion Research, 14, 247-263.

Tiedge, J. T., Silverblatt, A., Havice, M. J. & Rosenfeld, R. (1991). Discrepancy between perceived first-person and perceived third-person mass media effects. Journalism Quarterly, 68, 141-154.

Trope, Y. (1989) Stereotypes and dispositional judgment. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski and W. Stroebe (Eds.), Stereotyping and Prejudice: Changing Concepts (pp. 133-149). New York: Springer-Verlag.

Weiner, B. (1990). Searching for the roots of applied attribution theory. In S. Graham & V. S. Folkes (Eds.), Attribution theory: Applications to achievement, mental health, and interpersonal conflict (pp. 1-13). Hillsdale, NJ: Lawrence Erlbaum Associates.

White, H. A. (1997). Considering interacting factors in the third-person effect: Argument strength and social distance. Journalism and Mass Communication Quarterly, 74, 557-564.

Wu, W. & Koo, S. H. (2001). Perceived effects of sexually explicit internet content: The third-person effect in Singapore. Journalism and Mass Communication Quarterly, 78, 260-274.

Youn, S., Wan, F. & Faber, R. J. (2001). We should censor because they are vulnerable: Censorship of controversial web sites and the third-person perception. Proceedings of the 2001 Conference of the American Academy of Advertising, C. R. Taylor (ed.), 72-81.

Table 1

Mean Scores for Recent High School Graduates in Assessing Influence of Advertising

Ad	**Self**	**Ad Mjrs**	**Mid-40s**	**Mid-70s**
Depend	1.33	3.13	3.95	8.33
Lincoln	4.50	5.74	7.34	5.38**
Propel	5.83	6.25*	4.09	1.96
Remdex	4.01	5.14	6.29	6.53

Table 2

Mean Scores for Senior Ad Majors in Assessing Influence of Advertising

Ad	**Self**	**HS grads**	**Mid-40s**	**Mid-70s**
Depend	1.12	.53	4.19	8.38
Lincoln	4.69	2.53	7.56	6.21
Propel	5.94	7.11	4.07	1.82
Remdex	4.79	3.43	6.65	6.61

On both tables mean scores are on a 0 to 10 scale of influence with 0 for "not at all" and 10 for "a great deal. Comparisons are based on t-tests between the mean for "self" and other means in the same row. Differences in means are significant at $p< .0001$ unless otherwise indicated. * $p < .05$; ** $p < .001$

Table 3

Mean Score Comparisons--Recent High School Graduates and Senior Advertising Majors

Ad and Group Assessed	Recent HS Mean	s.d.	Ad Major Mean	s.d.
Depend H.S.	1.33**	2.74	.53	1.05
Depend Ad Mjrs	3.13**	2.74	1.12	1.73
Depend mid-40s	3.95	2.32	4.19	2.04
Depend mid-70s	8.33	2.10	8.38	1.68
Linc. H.S.	4.50**	2.55	2.53	2.03
Linc. Ad Majors	5.74**	2.53	4.69	2.29
Linc. Mid-40s	7.34	1.98	7.56	1.52
Linc. Mid-70s	5.38**	2.52	6.21	2.09
Propel H.S.	5.83**	2.43	7.11	1.90
Prop. Ad Mjrs	6.25	2.27	5.94	2.17
Propel Mid-40s	4.09	2.11	4.07	1.73
Propel Mid-70s	1.96	1.97	1.82	1.57
Remdex H.S.	4.01*	2.58	3.43	2.14
Rem. Ad Mjrs	5.14	2.59	4.79	2.23
Rem. Mid-40s	6.29	2.02	6.65	1.85
Rem. Mid-70s	6.53	2.24	6.61	2.26

(0 = no influence; 10 = "a great deal.")

Differences in means as determined by t-tests are not significant unless indicated.
* $p < .05$; ** $p < .001$

Table 4

Attributions of Recent High School Graduates

You have just made evaluations as to how advertising might influence different groups of people. What ideas/experience/information, etc., did you draw up to estimate the influence of advertising on:

Senior Advertising Majors? (n = 144)

52.8% Mentioned aspects of age group
3.6% Said they knew or know people of that age
52.8% Mentioned product and/or ad or targeting to age group

People in Their Mid-40s? (n = 140)

50.0% Aspects of age group
11.4% Knew or know people of that age
48.6% Product and/or ad or targeting to age group

People in their 70s? (n = 141)

51.1% Aspects of age group
9.2% Knew or know people of that age
51.4% Product and/or ad or targeting to age group

Table 5

Attributions of University Seniors Who Are Advertising Majors

You have just made evaluations as to how advertising might influence different groups of people. What ideas/experience/information, etc., did you draw up to estimate the influence of advertising on:

Recent High School Graduates? (n = 140)

46.4% Mentioned aspects of age group
14.3% Said they knew or know people of that age
64.3% Mentioned product and/or ad or targeting to age group

People in Their Mid-40s? (n = 136)

55.2% Aspects of age group
12.5% Knew or know people of that age
56.6% Product and/or ad or targeting to age group

People in their 70s? (n = 138)

58.0% Aspects of age group
8.7% Knew or know people of that age
51.4% Product and/or ad or targeting to age group

Modeling micro and macro: A multilevel model to predict memory for television content

Brian G. Southwell*

Paper presented at the 2003 AEJMC convention (Kansas City, Missouri, USA)

Communication Theory & Methodology Division

*Southwell is an Assistant Professor at the University of Minnesota.

Author contact information

Mailing address:
Dr. Brian Southwell
School of Journalism and Mass Communication
111 Murphy Hall
206 Church Street SE
Minneapolis, MN 55455

E-mail: south026@umn.edu
Phone: 612-624-2491

AUTHOR'S NOTE: Data reported in this paper result partially from work funded by the National Institute on Drug Abuse (contract number N01DA-8-5063) through a primary contract with Westat, Inc., of Rockville, Maryland, and a subcontract with the University of Pennsylvania's Annenberg School for Communication. I also am grateful to Ogilvy & Mather of New York for providing the media time purchase data used in the analyses. Robert Hornik of the University of Pennsylvania and David Maklan of Westat are Co-Principal Investigators for the NIDA project.

Modeling micro and macro: A multilevel model to predict memory for television content

Abstract

Whenever a study engages an array of variables that should involve different units of analysis, the risk of misleading results lurks. Questions about memory for media content, for example, invite investigation of not only variables describing individuals, but also (relatively speaking) macro-level constructs concerning content. This paper uses multilevel modeling techniques to avoid basic pitfalls and predict memory for electronic media content using data from U.S. adolescents and data regarding nationally available health campaign advertisements.

Modeling micro and macro: A multilevel model to predict memory for television content

In a crucial issue of *Communication Research* on units of analysis more than 10 years ago, Price, Ritchie, and Eulau (1991) argued that much communication research lies at an intersection of macro-level theorizing and available micro-level measurement and could be informed by cross-level or multi-level approaches. That same issue also included Pan and McLeod's (1991) recommendation to let the theoretical locus of variance of the dependent variable in question and the mechanisms hypothesized to account for that variance determine the appropriate level of analysis. Nevertheless, truly multilevel approaches are not yet widespread or common in communication research in general. We return to that call here with an effort to use multilevel modeling techniques to predict memory for content from a recent national strategic communication campaign.

Given the role of exposure as an explanation for the presence or lack of campaign effects (Hornik, 1997), locating and understanding the immediate imprint of exposure in individuals is a worthwhile endeavor. If one asks what predicts memory for exposure to media-based campaign efforts, in turn, there are a variety of candidate explanations that arise. As we might expect in light of the discussion above, those explanations do not all reside on the same plane: individual-level explanations, for example, contrast with explanations that concern some aspect of the media content in question.

On one level, Cappella (1996) has argued that investigation about possible media effects should begin with consideration of the individual mental processes and structures that constrain audience member responses. Studying media exposure among humans, after all, means that biological and cognitive constraints bound what is possible. While such individual-level consideration is undoubtedly relevant and useful, nonetheless, all individual engagement with electronic media also occurs in a social, cultural, institutional, and organizational context (Pan & McLeod, 1991; Wright, 1986). Certainly, for example, the simple environmental prevalence of particular media content should affect individual memory for it in some fashion. What is most appropriate, then, is to understand memory for media content as the likely product of a multilevel model of predictors.

Whenever a study, such as the present one, engages a series of variables that by definition should be located on different planes of measurement, the risks of misdirected assignment of units of analysis and misleading results lurk (Bryk & Raudenbush, 1988; Burstein, 1980; Haney, 1980). Bryk and Raudenbush (1988), for example, point out that education data are routinely analyzed solely at the student level. Such a move assumes that educational interventions or organizational contexts, i.e., school-level variables, are constant across all students. Insofar as effects vary both among students and among contexts, conventional approaches may be misleading. Similarly, all media campaign content is not equal, either in terms of general environmental prevalence or in terms of various content features. In light of those ideas, we will outline and explicitly assess here a multilevel model of individual memory for exposure, or what we can call encoded exposure, to a recent national campaign effort.

Encoded exposure as a dependent variable

Theoretically, an individual can be said to have encoded exposure for any particular unit of media content when he or she holds a retrievable memory trace (available upon prompting) that corresponds to that content and offers some sense of the frequency of past engagement with that content. Undoubtedly, researchers have identified a plethora of existing memory systems and types that differ considerably in complexity and nature in comparison to this basic construct (see Bower, 2000, for a discussion). Nonetheless, while other aspects of memory also are noteworthy, this basic concept should be central for the purposes of many campaign evaluations and offers a reasonable focal point for the present study.

Given the notion of a minimal memory trace, then, at least two individual memory performance task options are relevant as potential measures: a recognition task or a recall task.

The two types of memory measures are related; measures of each often covary (Singh, Rothschild, & Churchill, 1988; Zinkhan, Locander, & Leigh, 1986). Nevertheless, recognition can be differentiated from unaided recall of information. We can think about unaided recall as the ability to offer detail about particular content when asked an open-ended question at some point after initial opportunity to engage the content. Recognition, in contrast, is a more basic ability to respond to a closed-ended question about past engagement with specific content when presented that content once again. Whereas recall suggests a relatively high degree of current information salience and accessibility, recognition involves a somewhat lower standard of past cognitive engagement (Shoemaker, Schooler, & Danielson, 1989; Singh et al., 1988).

In light of this distinction, recognition-based tasks theoretically should offer appropriate indicators of encoded exposure. As Lang (1995) has argued, recognition measures likely indicate if the information in question ever has been encoded, suggesting that such encoding resides at a different conceptual level than the retrieval ability likely tapped by recall tasks. While unaided questions may provide a keener sense of what is most salient to a respondent at the time of interview, measuring recognition should more precisely and efficiently tap basic encoded exposure (du Plessis, 1994; Stapel, 1998). This paper describes such a recognition-based measure, validates that measure through a demonstrated relationship between it and the simple environmental prevalence of media content, and then explores additional predictors of the variable through multilevel modeling efforts.

Hypothesized individual-level and content-level predictors

There are several individual-level variables that should matter in predicting encoded exposure for health advertisements that appear on television. On a simple level, these variables must include individual television use and other indicators of opportunities for exposure. More extensive television watching should lead to higher levels of encoded exposure.

The relevance of an advertisement's topic also should matter. The more cause to believe that a particular type of media content should be relevant for an individual, the more likely that individual should be to report encoded exposure, all else being equal. Two prominent and complementary social psychological models of persuasion, namely the Elaboration Likelihood Model (Petty & Cacioppo, 1986a; Petty & Cacioppo, 1986b; Petty & Priester, 1994) and the Heuristic-Systematic Model (Chaiken et al., 1989; Eagly & Chaiken, 1993), offer some relevant insight in this regard. Though minor differences[1] can be enumerated, the two models converge to suggest that perceived personal relevance motivates effortful processing and should lead to encoded exposure (by virtue of affecting depth of processing and facilitating storage in memory).

Increased perception of the personal relevance of a message is associated with increased thinking about that message (Brickner, Harkins, & Ostrom, 1986; Leippe & Elkin, 1987; Petty, Cacioppo, & Haugtvedt, 1992). Increased elaboration, in turn, should be predictive of more enduring possibility for later retrieval or recognition of the various instances in which a message was encountered in one's media environment. Variables indicating ostensible personal relevance of particular media content, then, should positively affect an individual's encoded exposure to that content. (With regards to anti-drug advertisements, the focus of the present study, a key indicator of perceived relevance will be the extent of one's past drug use.)

In addition, conversation with others about the general topic of the advertisement also should bear a relationship (at least one of association if not causation) to later reports of encoded exposure. Engagement with mass media does not occur in a vacuum. Social networks play a role in shaping a person's initial engagement with such content, their retention of such engagement, and their action as a result of such engagement (Hagen & Wasko, 2000; Hornik, 1989; Katz & Lazarsfeld, 1955; Wright, 1986). Accordingly, the degree to which someone has conversations with others about the general topic of the content in question also should predict reported exposure in a positive manner.

There are two ways in which conversations that do not necessarily explicitly refer to particular television content could nonetheless impact encoded exposure reporting about that content. First, in the present case of anti-drug advertisements, a person who has engaged such an advertisement and who then discusses the general topic of drugs with another person might reinforce their cognitive imprint of the content in question through activation of related nodes during the course of conversation. Theoretical backing for this idea lies in Anderson's (1983; 1990) network model of memory, which both posits the possibility that repeated activation of certain memory nodes can reinforce the accessibility of adjacent nodes.

Insofar as information units related to "marijuana" are stored in connected memory nodes that are activated every time a person encounters the word, for example, conversation about drugs should arouse or activate not only nodes directly involved in that conversation, but also nodes where images of anti-drug advertisements are stored. In this manner, conversation about the topic should make any stored image of anti-drug advertising more salient and should increase the likelihood of that person recognizing the advertisement when it is presented in a survey.

A second possibility is that conversation about drugs provides cognitive fodder for later processing and recognition of related media content. A person who has a conversation with another person about drugs in general might bolster or enrich their schemata with reference to drugs such that they later engage a particular presentation of drug-related media content more efficiently than they would have otherwise. In turn, they should be more likely to report encoded exposure for unit of media content when presented with it in the future.

At the advertisement content level, both the sheer prevalence of an advertisement and the formal features of that advertisement should predict (average) encoded exposure. The justification for including a prevalence variable in our model is straightforward. Certainly, the simple environmental prevalence of particular media content should affect individual exposure to it in some fashion; without such prevalence, we could not hope for widespread memory of past engagement. Many commercial entities underscore this point, depicting exposure, for example, as a function of simple correspondence between the prevalence of content within an information environment and aggregate availability of individuals to engage that content (Webster, Phalen, & Lichty, 2000).

Beyond the simple environmental prevalence variable noted above, at least one formal feature of advertisements should affect the degree to which respondents report encoded exposure. Specifically, what we can call the context instability of a unit of media content should bear a generally negative relationship to reported encoded exposure for that media content. Context instability refers to the degree to which a unit of media content transitions between distinct depictions of time or space, transitions that in combination should present significant processing hurdles for individuals.

Evidence and arguments from a variety of sources highlight the relevance of depicted transitions between different points in either time or space that transcend normal human expectations for movement through either of those dimensions. For example, the limited-capacity approach to understanding human engagement with mass media, which builds upon earlier work by Broadbent (1958) and has been posited cogently by Lang (2000), suggests that individuals are limited in their ability to process media content by cognitive capacity constraints. The approach suggests that content sometimes can overload one's processing system, resulting in presented information not being processed and stored.

At the center of this potential for overload is the frequency of new information appearance and the processing it demands. While information-rich presentations can arouse attention under some circumstances, Lang and others (Lang, 1995; Lang, 2000; Lang, Geiger, Strickwerda, & Sumner, 1993) have suggested that formal features of a message that introduce substantial amounts of new information also can inhibit processing and later recognition ability. Visual context instability, then, should affect the memory encoding potential for media content insofar as it tends to overtax individual processing systems. The greater the context instability presented, the less encoded exposure we should expect, all else being equal.

Justification of a formal multilevel approach

By separating analyses into individual-level and advertisement-level approaches, we could present initial evidence that encoded exposure is rightly understood as a product of multiple levels of predictors. At the same time, research on multilevel modeling, e.g., Rowan, Raudenbush, & Kang (1991) and Sampson, Raudenbush, & Earls (1997), suggests that simultaneous estimation of all predictor levels is more appropriate. Also, there are additional worthwhile analyses to explore. Formal fitting of a multilevel model will highlight answers to three important questions about memory for campaign advertisements among adolescents: one regarding the (hypothesized) multilevel distribution of encoded exposure variance, one regarding the plausibility of a multilevel model, and one regarding possible cross-level interactions.

Any theory positing that encoded exposure to media content warrants a multilevel understanding assumes that a content-level grouping of data generated to study the phenomenon will account for a significant amount of the overall variance in the dependent variable. To test that assumption here, initial assessment of the intraclass correlation[2] as it relates to a specific advertisement grouping will offer a sense of the specific proportion of total variance in encoded exposure that lies between advertisements.

Beyond data structure questions, do the various predictors hypothesized above demonstrate significance when included in a single multilevel model? To address this question, we need an approach that affords explicit modeling at two levels of analysis so that the estimated effects of independent variables at one level of analysis can be adjusted simultaneously for effects at the other level of analysis. Such an analysis is presented here.

Lastly, in addition to main effects, advertisement-level predictors may curtail or attenuate the effects of independent-level variables on encoded exposure. Fitting a multilevel model will shed light on whether that is the case. We can assess whether a significant amount of random variation exists in any estimated individual-level predictor coefficient associated with initial model estimation. Such random variation in a coefficient is predictable (potentially) as a function of advertisement-level variables. A multilevel approach not only estimates individual-level effects within each macro-level group but also assumes that such individual-level effects might vary between groups as a function of macro-level variables. For any such compelling possibilities, we can model the individual-level coefficient in question as a function of content-level predictors, i.e., environmental prevalence and context instability.

Methods

Procedure

Beginning in 1999, the National Survey of Parents and Youth (NSPY) has been funded by the National Institute on Drug Abuse to evaluate federal government efforts to discourage drug use through a national media campaign. As a part of those media-based efforts, campaign organizations placed anti-drug advertisements in national network, cable, and in-school television programming, as well as in local television programming in over 100 U.S. metropolitan areas. One of the main objectives of NSPY is to track memory for, and assess the impact of, those advertisements among U.S. adolescents.

From November 1999 through December 2000, a multistage cluster sample[3] representing all U.S. youth ages 9- to 18-years-old and their parents or caregivers participated in two waves of NSPY. In a first wave, from November 1999 through May 2000, interviewers administered surveys with 3,312 youth aged 9 to 18 in 2,373 households. From July 2000 through December 2000, interviews also were conducted with 2,362 youth aged 9 to 18 in 1,726 households. Respondents used touch-screen laptop computers and headphones brought into their homes by an interviewer to view each question (or listen to a prerecorded reading of the question) and to respond. For a complete discussion of the first two waves of the NSPY study, see Hornik et al. (2000) and Hornik et al. (2001).

The first challenge to be met in fitting a multilevel model to NPSY data was organizational in nature. More than 5,000 adolescents contributed responses for the two waves of NSPY analyzed. Each respondent contributed data in response to a series of interview presentations involving up to four advertisements from the 23 general market advertisements from the campaign (as discussed in detail below in the measures section). This situation resulted in a stacked dataset, whereby each respondent contributed more than one case of advertisement-specific measures.

In order to organize that data into usable form for a multilevel modeling endeavor, several steps proved useful. First, all cases corresponding to either non-eligible or non-general-market advertisements were removed from the dataset. For example, cases involving bogus advertisements that were shown to NSPY respondents but that did not actually air were removed. Second, one case was selected randomly from each respondent.[4] This move resulted in an initial set of 5,521 cases. After sorting this data by the name of the advertisement for which a respondent contributed data, advertisement-level variables for the 23 advertisements then were merged and linked to the 5,521 cases.

From this original set of 5,521, 9- to 11-year-old respondents and others with missing values on the main independent variables (reiterated below) were dropped via listwise deletion from analyses for this study. The default dataset of 12- to 18-year-old respondents for all analyses in this paper has an n of 2,623. The resulting data set allowed analysis of both individuals and of 23 groups of individuals (grouped by advertisement).[5]

Measures

Dependent variable measurement warrants special attention, given the multilevel nature of the present challenge. Fortunately, NSPY included a variety of questions that afford appropriate measurement of encoded exposure. During each NSPY interview, campaign television advertisements that had aired in the two months prior to a particular interview were shown to respondents on the laptop computer used for the interview. Generally, the interview program played up to four advertisements for respondents, depending on the number of eligible advertisements. After seeing each advertisement, each respondent was asked, "Have you ever seen or heard this ad?" If they responded in the affirmative, they then were asked, "In recent months, how many times have you seen or heard this ad?" Response categories were "not at all," "once," "2 to 4 times," "5 to 10 times," and "more than 10 times." In order to produce a reasonable interval measure, these categories were recoded into scores of 0, 1, 3, 7.5, and 12.5 for analysis. "Don't know" responses to the initial question were recoded as 0.5. Summed across the general market advertisements eligible for a respondent, this recoded question offered an indicator of individual exposure.

At the individual level, we could assess encoded exposure across the series of advertisements shown to a person during an interview. Given the reorganized, multilevel dataset discussed above in the procedure section, however, it was more useful to look at the number of times a respondent reported being exposed to one randomly selected advertisement. This encoded exposure measure offers both individual-level variation, i.e., person-to-person variance, and aggregate-level variation, i.e., differences in mean levels of the measure between different advertisements. Such variation affords the basis for the multilevel analysis presented here: A single encoded exposure measure (EXPOSEAD) stands to be analyzed at two different levels simultaneously in the same multilevel model.

Independent variable measurement also warrants explanation. Four television use measures (TVUSE, TVPROGS, CABLE, and ONE), a past drug use indicator (LNUSEDEP), a measure of recent school attendance (MISSCHL), and at least one conversation variable (DRUGCONV) served as available indicators for the individual-level variables noted earlier in our discussion. A brief discussion of each follows below.

Several different NSPY questions in combination offered independent measures of various dimensions of television use. For example, all youths were asked, "How much TV do you estimate watching on an average weekday?" and were offered response categories including "none," "half-hour or less," six separate options for one through six hours, and "7 or more hours." Following that question, youths also were asked for an estimate of their TV watching during an "average weekend" and were offered categories including "none," "less than one hour," options for "1 to 2 hours" through "9 to 10 hours" and "11 or more hours." I combined responses from these two questions into a weekly estimate of television watching (TVUSE) by assigning interval-level numbers to each of the categories[6], multiplying the weekday measure by five, and adding the weekday total to the weekend measure.

In addition, for 12- to 18-year-olds, NSPY also included up to 15 questions regarding whether the respondent had ever seen particular television shows. Shows included in each wave of surveys were selected from the list of primetime and daytime shows (including both general market and highly watched African-American shows) in which national anti-drug campaign staff intended to purchase airtime, such as "ER," "Dawson's Creek," and "The Steve Harvey Show". Respondents who read (or listened to) and answered the survey exclusively in Spanish were presented with a list of Spanish-language shows targeted by the campaign. As a result, this measure also offered an indicator of a respondent's opportunity for engagement with campaign advertisements by virtue of their engagement with relevant television content. For analysis purposes, all of the items were dichotomized into two categories: having "never" seen a show or reporting at least some past watching. The items then were combined into an additive index (TVPROGS) that ranged from zero to 15.

Because the ONDCP campaign focused not only on network television, which is largely available to most American youths, but also on venues such as cable television and in-school programs such as Channel One, two additional measures of television use also are useful. In reference to cable programming, 12- to 18-year-old respondents were asked how often in the past 30 days had they watched different types of channels: channels focused on music television, all-sports programming channels, channels with programming intended primarily for African

Americans, or Spanish-language channels (for those interviewed in Spanish). After converting original response categories into interval levels[7], these measures were added together to construct an index of relevant cable programming use (CABLE). In regards to in-school programming, a NSPY question asked of 12- to 18-year-olds regarding drug-related information available via Channel One includes the option to report that one's school does not have the channel. This measure afforded a dichotomous indicator of Channel One use (ONE). Tendency to miss class (MISSCHL) was measured with a question asking how many days in the past 30 days one had skipped school. (Because it serves as a simple indicator of opportunity for in-school exposure, MISSCHL should bear a negative relationship to EXPOSURE.)

USEDEPTH indicates the depth of an adolescent's past marijuana use, depending on whether a respondent reported no past marijuana use whatsoever (USEDEPTH = "1"), previous trial but no regular use (USEDEPTH = "2"), or any previous instance of regular use (USEDEPTH = "3"). (Because of skewness in the USEDEPTH distribution, analysis presented below employs the natural log of the measure, which we can call LNUSEDEP.[8])

With regard to conversation about drugs, all youth NSPY respondents were asked, "In the last 6 months, how often have you and either of your {parents/caregivers} talked about drugs?" Available response categories included "Never," "Once," "2 to 3 times," "4 to 5 times," "6 to 10 times," and "More than 10 times". Similarly, youth respondents were asked, "In the last 6 months, how often have you and your friends talked about drugs?" Similar response categories were offered. For both questions, a recoded[9] measure offered an interval-level indicator of recent drug conversation frequency. The analyses presented here employs a single summary measure of drug conversations (DRUGCONV) that is a simple additive index combining frequency of recent drug conversations with parents or caregivers and frequency of such conversations with friends.

In addition, because the advertisements in question vary with regard to the age and race or ethnicity of people depicted, dummy indicators of race and ethnic groups (AFAM, HISP, and OTHER, in comparison to WHITE as a reference group) and age (D14to15 and D16to18, in comparison to 12- to 13-year-olds as a reference group) also were included in the model presented and relevant interactions were explored.

At the content level, sources beyond NSPY provided measures of environmental prevalence (GRPS) and context instability (LNCUTS). For example, a Gross Rating Points (GRPs) estimate for each advertisement, as reported by campaign contractors based on estimates of the reach and frequency obtained for each advertisement, served as a reasonable proxy for the environmental prevalence of a particular advertisement. A GRP is a conventional unit used by advertisers to measure a population's simple physical opportunities for exposure to media content and is the product of underlying estimates of reach and frequency (Farris & Parry, 1991).

Measuring context instability, or the degree to which a unit of media content depicts different locations in time or space in sequence, offers an additional challenge. The present study uses as a measure of context instability the number of cuts per second in a campaign advertisement.[10] Insofar as a cut here is essentially a transition from one depicted location in time or space to another, that operational definition should offer a useful measure of the construct described earlier.

Analysis

The family of multilevel models known as hierarchical linear models (Bryk & Raudenbush, 1992) offers a reasonable set of tools for the present challenge. Estimation of a hierarchical linear model (HLM) often is more appropriate than ordinary least squares regression (OLS) methods because HLM acknowledges a unique error structure at each level, whereas OLS approaches do not automatically do so (Bryk & Raudenbush, 1988; Bryk, Raudenbush, Congdon, & Seltzer, 1986). Such models have been applied to a variety of research problems, including modeling academic achievement as a function of student and school variables, e.g., Rowan, Raudenbush, & Kang (1991), and understanding individual and neighborhood crime variables, e.g., Sampson, Raudenbush, & Earls (1997). We also should be able to apply them here.

Accordingly, version 5.03 of the HLM program (Raudenbush, Bryk, & Congdon, 2001), which offers maximum likelihood estimation of hierarchical linear models, was useful for this study.

The HLM framework directly accommodates the three major issues posed earlier. The question of whether a multilevel model is more appropriate than a single-level model, for example, can be addressed by looking at two types of statistics: intraclass correlation and reliability estimate of group means. Careful explication of the basic equations underlying these statistics will facilitate all later discussion and so is quite worthwhile.

HLM 5 allows assessment of the degree to which dependent variable variance can be decomposed into significant within-group, e.g., individual-level, and between-group, e.g., advertisement-level, components. Two equations, adapted from Rowan, Raudenbush, & Kang (1991), illustrate this decomposition.

1) Within-advertisement-group model

$$Y_{ij} = \beta_{0j} + r_{ij}$$

Y_{ij} is the encoded exposure score for respondent i in advertisement group j, β_{0j} is the mean score for the advertisement group, and r_{ij} is a random error for individual i in group j that is normally distributed with mean 0 and variance σ^2. The within-group variance (σ^2) will prove useful below.

2) Between-advertisement-group model

$$\beta_{0j} = \nu_0 + U_{0j}$$

In this equation, ν_0 is the grand mean of encoded exposure and U_{0j} is a random error term that is normally distributed with mean 0 and variance τ.

These two equations parallel a standard one-way random effects ANOVA model for this situation, in which advertisement group would be considered to be a random factor with varying numbers of respondents in each group. Following from these two equations, we can use the within- and between-group variance components to compute an intraclass correlation with the following equation, also adapted from Rowan, Raudenbush, & Kang (1991).

3) Intraclass correlation

$$\rho = \tau / (\sigma^2 + \tau)$$

In this instance, the ρ parameter essentially is an estimate of the proportion of total variance in encoded exposure that lies between advertisement groups. A relatively high ρ value would suggest that a relatively large amount of the total variance in encoded exposure lies between advertisements. If a sizable amount of variance can be classified as lying between advertisements, then we will have further evidence of the necessity of approaching encoded exposure as a function of multilevel influences.

Based on these components and the sample size of each group, HLM also offers easy calculation of a measure of the reliability of an estimated group mean. For each group, HLM computes a reliability estimate, α_j, with the equation, $\alpha_j = \tau / (\tau + \sigma^2/n_j)$, where n_j is the sample size for group j. We then can assess the average reliability of the advertisement group mean by looking at the value of α_j / k, where k is the number of advertisement groups (23, in the present analysis). If the average reliability for all groups is relatively high, then we also can have further confidence that between-group analyses of encoded exposure can be presented with relatively less concern about potential dependent measure error (Rowan, Raudenbush, & Kang, 1991).

Answers to both the second and third research problems posed above also can draw upon HLM results as useful evidence. Before addressing complex questions of cross-level interactions, for example, it is crucial to know first whether a simultaneously estimated two-level model of encoded exposure composed of hypothesized predictors lends any support to our speculation about main effects. For this purpose, the HLM 5 program allows simultaneous estimation of the following two equations (using restricted maximum likelihood methods to generate parameter estimates and robust standard errors for those estimates).[11]

4) Level one model

$\mathrm{EXPOSEAD} = \beta_0 + \beta_1\,(\mathrm{TVUSE}) + \beta_2\,(\mathrm{TVPROGS}) + \beta_3\,(\mathrm{CABLE}) + \beta_4\,(\mathrm{ONE}) + \beta_5$
$(\mathrm{AFAM}) + \beta_6\,(\mathrm{HISP}) + \beta_7\,(\mathrm{OTHER}) + \beta_8\,(\mathrm{LNUSEDEP}) + \beta_9\,(\mathrm{DRUGCONV}) + \beta_{10}$
$(\mathrm{D14to15}) + \beta_{11}\,(\mathrm{D16to18}) + \beta_{12}\,(\mathrm{MISSCHL}) + r$

5) Level two model

$\beta_0 = v_{00} + v_{01}\,(\mathrm{GRPS}) + v_{02}\,(\mathrm{LNCUTS}) + u_0$

Also, each predictor coefficient is considered to be a function of an intercept and error term. For example, $\beta_1 = v_{10} + u_1$.

Beyond these parameter estimations, we also will want to talk about the degree to which any estimated overall model explains variance in encoded exposure. A useful and computable statistic for this purpose is the proportion reduction arising from the introduction of an explanatory model (relative to the simple two-level model without predictor variables outlined in equations 1 and 2). This proportion reduction can be interpreted as an indicator of the strength of the explanatory model and can be calculated separately for each level of a proposed two-level model (Bryk and Raudenbush, 1988). Individual-level and advertisement-level explanatory power, in this framework, can be assessed with the following equations.

6) Proportion variance reduction for level one

$$\frac{(\sigma^2 \text{ of model 1}) - (\sigma^2 \text{ of model 2})}{(\sigma^2 \text{ of model 1})}$$

7) Proportion variance reduction for level two

$$\frac{(\tau \text{ of model 1}) - (\tau \text{ of model 2})}{(\tau \text{ of model 1})}$$

In addition to producing fixed effects estimates to support or overturn hypothesized relationships, the HLM program also estimates residual variance components for all of the individual-level predictor slopes estimated. This information will shed light on the third issue raised earlier, namely the possibility of cross-level interactions. Indications of a significant amount of residual variance remaining in the estimated slope for a first-level predictor will suggest the potential usefulness of a more extensive model that includes slopes as outcomes.

In such a more elaborate model, second-level predictors would not only account for differences in group means but also can account for differences in first-level predictor slopes. Not only β_0 but also β_1, for example, might be a function of content prevalence or content features. In that instance, HLM can produce estimates for the following model: $\beta_1 = \nu_{10} + \nu_{11}$ (GRPS) + ν_{12} (LNCUTS) + u_1. When appropriate, we also can test such additional models below.

Results

Within-advertisement-group versus between-advertisement-group variance

Decomposition of the variance in EXPOSEAD suggests that a significant and sizable proportion of the variance lies between advertisements, $t = 5.14$, $df = 22$, $p < .01$. Drawing upon equation 3 from above and the estimated values of $\sigma^2 = 11.07$ and $\tau = 1.75$, we can see that $\rho = .14$. This intraclass correlation suggests that approximately 14 percent of the total variance in encoded exposure lies between advertisement groups. In addition, the average reliability estimate for advertisement-group exposure means was 0.91, which justifies dependent variable measurement at the group level. Both findings suggest macro-level influence on memory.

Multilevel model of encoded exposure: Main effects

Table 1.1 summarizes the results of an estimated multilevel model. Both individual- and advertisement-level explanatory variables were successful in explaining variance in this context. The extent to which an adolescent had seen television programming targeted by the campaign, attendance at a Channel One school, and reported conversations about drugs all bear positive relationships to encoded exposure, $p < .01$ for each. In addition, past drug use bears a negative relationship to encoded exposure, $p < .01$. In comparison to 12- to 13-year-olds, 16-to-18-year-old respondents report less encoded exposure and white respondents report more encoded exposure than respondents who are not African-American, Hispanic, or white. Moreover, GRPs predict encoded exposure in a positive fashion and context instability holds a negative relationship to the dependent variable, $p < .01$ for each, as predicted.

Relatively speaking, this model appears to account for a greater percentage of the explainable between-group variance in encoded exposure than of the within-group variance (though it is worthwhile to recall that the majority of overall exposure variance lies at the individual level in this sample). At the individual level, σ^2 initially was 11.07 and is 9.61 after estimation of this explanatory model, resulting in a 13 percent reduction of variance. At the advertisement level, τ initially was 1.75 and is 1.29 after estimation of this explanatory model, resulting in a 26 percent reduction of variance.

Table 1.1

Multilevel model of encoded exposure (equations 4 and 5)

Variable	B (predicting group mean)	B (mean fixed effect)	SE B	df
Level one (n = 2,623)				
TVUSE		.01	0.01	22
TVPROGS		.10**	0.02	22
CABLE		.01	0.003	22
ONE		.31**	0.11	22
Race/ethnicity				
African-American		.30	0.30	22
Hispanic		-.11	0.20	22
Other		-.77*	0.28	22
LNUSEDEP		-.04**	0.01	22
DRUGCONV		.07**	0.02	22
Age comparisons				
14- to 15-years-old		-.15	0.21	22
16- to 18-years-old		-.49**	0.16	22
MISSCHL		-.11	0.08	22
Level two (23 groups)				
GRPS	.04**		0.003	20
LNCUTS	-.35**		0.07	20
Constant	-1.35**		0.35	20

Note. Via level two, this model accounts for 26 percent of encoded exposure variance between groups and, via level one, 13 percent of variance within groups. The reference groups for racial and ethnic and age comparisons are whites and 12- to 13-year-old respondents, respectively.
* $p < .05$. ** $p < .01$. Robust standard errors are reported, as recommended by Raudenbush, Bryk, and Congdon (2001), though estimation of fixed effects without robust standard errors told a similar story.

Beyond such results, however, the non-significant coefficient for TVUSE warrants attention. Could it be that the relationship of TVUSE to EXPOSEAD is a function of content-level influences? For some advertisements, the relationship between TVUSE and EXPOSEAD might be weak enough to dilute the average reported relationship. For example, a cross-level interaction between GRPs and TVUSE could have produced the above pattern; without any prevalence, no amount of TVUSE will produce exposure. We turn to that possibility next.

Multilevel model of encoded exposure: Cross-level interactions

We can assess the aforementioned cross-level influence possibility by looking at whether there is significant random variation in the TVUSE slope that is potentially attributable to an advertisement-level variable. For example, if we assume that the TVUSE slope itself is a function of $v_{10} + u_1$, then we can assess whether u_1 significantly differs from zero. Table 1.2 highlights the final estimation of such error terms associated with the results in table 1.1.

Table 1.2

Random effects for individual-level predictors from table 1.1

Variable	Random effect variance component	χ^2	df[12]
TVUSE	.0006*	31.44	18
TVPROGS	.007	18.79	18
CABLE	.00006	13.06	18
ONE	.10	14.32	18
Race/ethnicity			
African-American	1.01**	45.24	18
Hispanic	.26	17.26	18
Other	.42	11.29	18
LNUSEDEP	.002	17.78	18
DRUGCONV	.004*	32.94	18
Age comparisons			
14- to 15-years-old	.49*	34.89	18
16- to 18-years-old	.20	16.45	18
MISSCHL	.03	17.83	18
Constant	1.29	23.30	16

Note. * $p < .05$. ** $p < .01$.

Among other results[13], analysis of variance components does point to the existence of a significant random effect for the TVUSE slope, $\chi^2 = 31.44$, $df = 18$, $p < .05$. This suggests that there remains between-group variance in the relationship of TVUSE and EXPOSEAD that we can attempt to model as a function of level-two predictors. Additionally, table 1.2 also suggests that significant (and potentially explainable) between-group variance exists in the relationship of DRUGCONV to EXPOSEAD.

The possibility that both of these individual-level patterns are a function of macro-level influences is theoretically interesting. Such evidence could highlight the primacy of campaign information prevalence in determining the relationship of individual-level variables to reported campaign exposure. Such evidence also could demonstrate the amplification or dampening effect of individual variables for content-level influences.

We can test these possibilities by estimating a model that is identical to the model outlined above except that it also assumes coefficients for TVNEWS and DRUGCONV to not only be a function of a constant and an error, but also a function of GRPs and LNCUTS. In other words, we can assess the usefulness of including $\beta_1 = \nu_{10} + \nu_{11}$ (GRPS) + ν_{12} (LNCUTS) + u_1 and $\beta_9 = \nu_{90} + \nu_{91}$ (GRPS) + ν_{92} (LNCUTS) + u_9 among the elements to be estimated, where β_1 is associated with the main effect of TVUSE and β_9 is associated with the main effect of DRUGCONV.

If either content-level variable, i.e., GRPs or LNCUTS, is useful in accounting for variance in the TVUSE slope, for example, then we would expect the successful level-two predictor to garner a significant coefficient, e.g., ν_{11} or ν_{12} from the equation above. We would expect a similar pattern if either GRPs or LNCUTS can account for variance in the DRUGCONV slope. In addition, the new model including these new terms should account for even more advertisement-level variance than the model outlined in table 1.1.

Table 1.3 outlines the results from estimation of this alternative explanatory model. Results again highlight the predictive power of TVPROGS, ONE, LNUSEDEP, and age and racial and ethnic comparisons, $p < .01$ for each. Cross-level dynamics are also now apparent.

Table 1.3

Multilevel model of encoded exposure (with cross-level interactions)

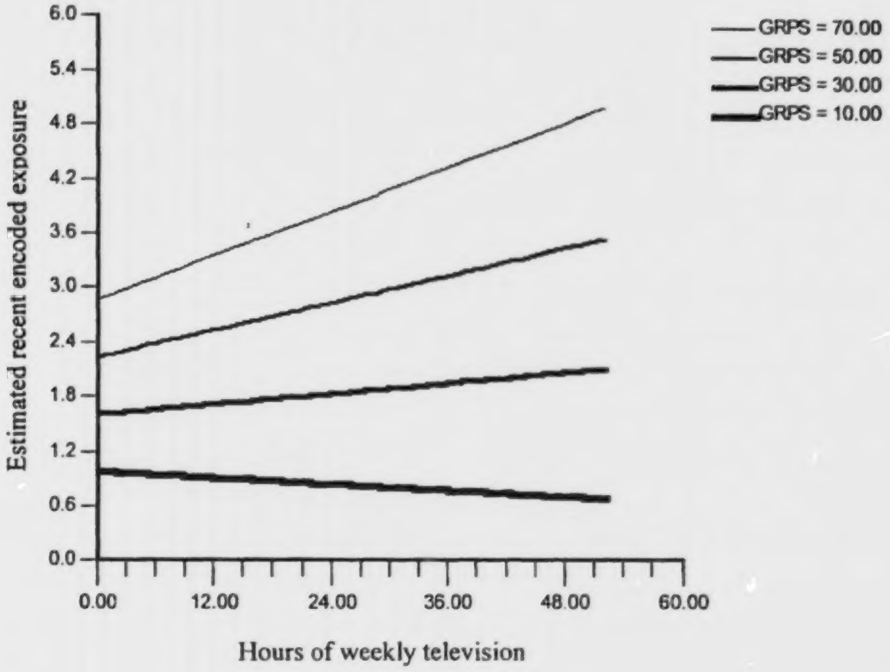
Variable	B	B	SE B	df
		(mean fixed effect)		
Level one (n = 2,623)				
TVUSE		-.02	0.01	20
TVPROGS		.10**	0.02	22
CABLE		.005	0.003	22
ONE		.33**	0.11	22
Race/ethnicity				
African-American		.32	0.30	22
Hispanic		-.10	0.20	22
Other		-.86**	0.27	22
LNUSEDEP		-.04**	0.01	22
DRUGCONV		-.01	0.02	20
Age comparisons				
14- to 15-years-old		-.13	0.22	22
16- to 18-years-old		-.50**	0.16	22
MISSCHL		-.12	0.08	22
Level two (23 groups)				
Prediction of level-one intercept				
GRPS	.02**		0.005	20
LNCUTS	-.24**		0.07	20
Constant	-.56		.34	20
Prediction of TVUSE B				
GRPS	.001**		0.0001	20
LNCUTS	-.002		0.002	20
Constant	-.02		0.01	20
Prediction of DRUGCONV B				
GRPS	.002**		0.0002	20
LNCUTS	-.02		0.01	20
Constant	-.01		0.02	20

Note. Via level two, this model accounts for 49 percent of the encoded exposure variance between groups and, via level one, 13 percent of the variance within groups. The reference groups for racial and ethnic and age comparisons are whites and 12- to 13-year-old respondents, respectively.
* $p < .05$. ** $p < .01$. Robust standard errors are reported, as recommended by Raudenbush, Bryk, and Congdon (2001), though estimation of fixed effects without robust standard errors told a similar story. (No probability of $p < .01$ reported above exceeded .05 in the non-robust analysis.)

The relationship between TVUSE and EXPOSEAD and the relationship between DRUGCONV and EXPOSEAD are associated with the environmental prevalence (GRPS) achieved by a particular advertisement. (LNCUTS is not a significant predictor in this capacity by conventional standards, though was marginally significant with regards to the DRUGCONV slope, p = .05.) In other words, the environmental prevalence of advertisements either moderates the relationship of particular individual-level variables or itself is moderated by such individual-level variables in its influence on encoded exposure. Television use, for example, appears to have a markedly different relationship with exposure depending on the degree to which the advertisement in question was prevalent on U.S. airwaves. Figure 1.1 illustrates this relationship.

Figure 1.1

Cross-level interaction (GRPs and TVUSE) to predict exposure



For campaign television advertisements that received prominent airplay, individual television use plays a significant role in explaining encoded exposure. For advertisements receiving little such airplay, however, individual television use is not an important predictor. We see an upward slope between TVUSE and EXPOSEAD at high levels of GRPs, whereas the relationship between TVUSE and EXPOSEAD is essentially flat at the lowest levels of GRPs.

A similar pattern exists with regard to the predictive ability of past conversation about drugs. As table 1.3 suggests, the positive relationship between DRUGCONV and EXPOSEAD is strongest for those advertisements for which campaign staff purchased or obtained a relatively high degree of environmental prevalence.

Importantly, inclusion of GRPs as a predictor of the relationship of TVUSE and DRUGCONV appears to have eliminated any significant random effects remaining for the coefficients of those two individual-level variables. While table 1.2 indicated significant variance in the coefficients initially estimated for each individual-level variable, the model fit and outlined in table 1.3 resulted in insignificant residual variance component estimates for TVUSE and DRUGCONV, $p > .10$ for each. This evidence again highlights the importance of paying attention to content-level prevalence differences.

Beyond these findings, however, we also can begin to parse out the directional nature of the conversation-exposure relationship. At least two possibilities are plausible. First, it might be the case that encoded exposure to anti-drug campaign advertisements (which itself is a function of environmental prevalence) simply tends to generate discussion, which explains the positive association between the two measures. As noted earlier, however, there are theoretical reasons to suspect a second possibility, as conversation about drugs might either sensitize a person's drug-related media content encoding tendencies or might arouse memory of past anti-drug advertisements and facilitate later recognition ability whenever drugs are discussed.

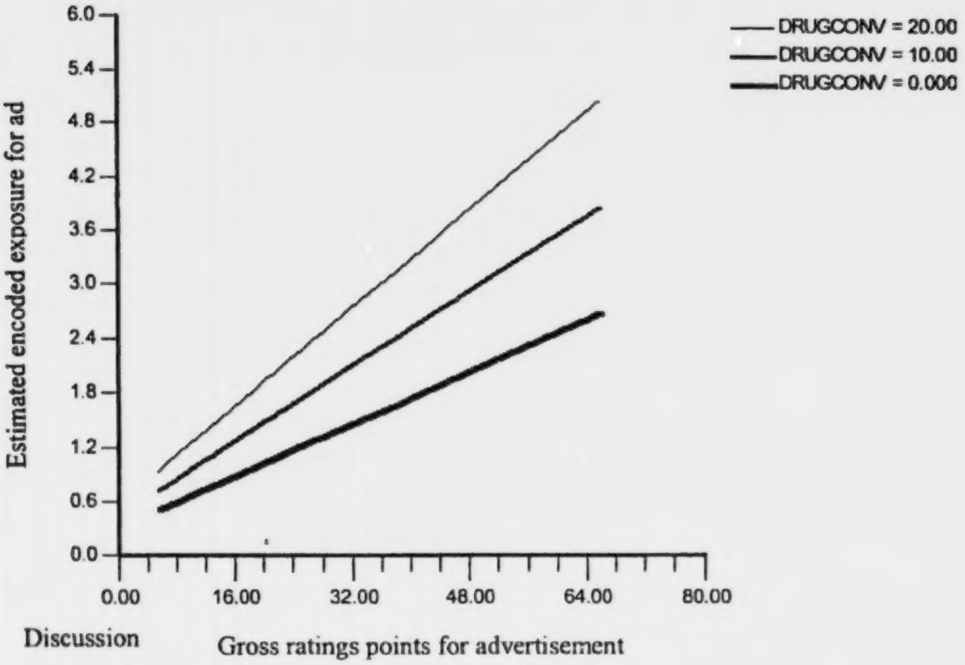
Results presented up to this point essentially go no further than demonstrating an association between conversation and encoded exposure and allowing for the reciprocal relationship possibilities. Because of the simultaneous estimation of both individual- and content-level effects presented, however, we also should be able to generate an additional piece of evidence regarding the nature of that conversation-exposure relationship by looking at the role of environmental prevalence. Specifically, we can ask whether widespread availability of media content leads to increased discussion or whether there is no relationship between macro-level anti-drug advertisement availability and micro-level discussion. In the first instance, we could view the individual-level conversation-exposure relationship as essentially a symptom of (or mechanism for) a general prevalence-conversation relationship. If there is no relationship between advertisement GRPs and the amount of drug conversation reported by respondents associated with that advertisement, however, then it will be reasonable to understand table 1.3 as suggesting that drug conversation moderates the impact of advertisement GRPs on encoded exposure. We might think of this phenomenon as a memory trace amplification effect.

Using DRUGCONV as a dependent variable, we can predict the mean level of drug conversation in advertisement respondent group simply as a function of GRPS and an error term. (This HLM analysis directly parallels the main analysis above in which GRPS predicted EXPOSEAD group mean). Results of this analysis undermine the possibility that reported general drug conversation is a function of the environmental prevalence of recent anti-drug advertisements. First, a decomposition of variance suggests that almost all of the variance in DRUGCONV lies within advertisement groups, not between them. Only roughly 1 percent (0.48 / 34.98) of the variance in DRUGCONV lies between advertisement groups. Second, GRPs do not bear a significant predictive relationship to the intercept of DRUGCONV, $\underline{B} = .007$, $\underline{SE}\ \underline{B} = 0.008$, $\underline{df} = 21$, $\underline{p} > .10$. These results suggest that conversations about drugs between adolescents and their friends and parents do not appear to be a function of the prevalence of specific campaign advertisements available during recent months.

In light of this pattern, general drug-related conversation in an adolescent's immediate social network (at least that network comprised of friends and parents or caregivers) appears to moderate the degree to which an anti-drug advertisement's prevalence translates into later memory trace retrieval. From this perspective, figure 1.2 depicts the cross-level interaction between GRPs and DRUGCONV in an appropriate manner, not only reiterating the general positive relationship between GRPs and encoded exposure but also suggesting that the relationship increases in strength when the number of drug conversation increases.

Figure 1.2

Cross-level interaction (GRPs and DRUGCONV) to predict exposure



Discussion

On an individual-level, then, encoded campaign exposure among 12- to 18-year-olds in the U.S. appears largely to be a function of their media habits, general conversation about drugs with friends and parents, and the extent of their own past drug use (albeit in a different manner than hypothesized with regard to the last predictor). Age and race differences also exist, which in part can be explained by targeting efforts on the part of ONDCP campaign staff.

The present results also offer some important contextual constraints for our discussion, however. For example, environmental prevalence and content features strongly predict encoded exposure levels; level two of the final model presented here accounts for about half of the group-level variance in encoded exposure. Nonetheless, it is also worth noting that total between-group variance represents a minority (about 14 percent) of the overall variance in encoded exposure among 12- to 18-year-old adolescents in the U.S., albeit a sizable minority. In other words, while we would be remiss to overlook macro-level effects when discussing encoded exposure (and in fact have avoided such an oversight here by documenting some striking macro-level effects), there is a considerable amount of individual-level variance that remains both outside the domain of macro-level main effects and unaccounted for by the individual variables highlighted here.

At the same time, the HLM efforts of the present study also offer more than simple confirmation or context. Allowing content-level variables not only to predict mean level of encoded exposure but also either to attenuate the relationships between individual-level variables and encoded exposure or to have their own relationships with exposure moderated by individual-level variables markedly improved the predictive power of the multilevel model in question. At the advertisement level, initial efforts accounted for approximately 26 percent of between-group variance in exposure, whereas an alternative model in which GRPs were allowed to predict the slopes of TVUSE and DRUGCONV in their relationships with exposure accounted for approximately one-half of all between-group variance in exposure. In other words, heeding the possibility for cross-level interaction resulted in a doubling of second-level predictive power.

If such an approach had not been taken, the cross-sectional nature of the individual-level measures employed in this study would limit discussion about the relationship between conversation and encoded exposure. In contrast, allowing a macro-level measure that theoretically precedes exposure encoding, i.e., environmental prevalence, to operate in a multilevel analysis afforded some clarification of the likely nature of the relationship between individual-level conversation and encoded exposure. Given the lack of a group-level relationship between GRPs and general drug conversation reported, past increases or decreases in advertisement prevalence do not appear to have preceded or (linearly) motivated recent general drug conversation involving 12- to 18-year-olds. Instead of solely being a product of encoded exposure, then, conversation, however it arises, appears to enhance memory retrieval ability for advertisements and also likely facilitates or moderates the tendency of an advertisement's environmental prevalence to result in encoded exposure reports. Without multilevel modeling results such as those highlighted here, such speculation would enjoy less empirical evidence.

Conclusions

By employing multilevel modeling techniques, this study produced three types of useful evidence regarding memory for television content among U.S. adolescents. First, basic variance decomposition confirmed that the distribution of encoded exposure itself invites a multilevel understanding. A significant and sizable proportion of exposure variance can be attributed to between-group differences when respondents are grouped according to the advertisement about which they were queried in the selected dataset. Second, an overall predictive model involving a variety of individual-level and content-level predictors confirms in most instances both the significance and the nature of the predictive power of each included variable. Beyond such results, the multilevel models fit presently also support the hypothesis that advertisement-level variables can interact with individual-level variables in having a joint effect on encoded exposure.

In general, then, the results highlighted here confirm that encoded exposure is rightly understood as a multilevel phenomenon. Importantly, however, this study also highlights ways in which multilevel modeling techniques, such as maximum likelihood estimation of hierarchical linear models, can be useful for approaching communication research questions involving both individual variables and variables that describe mass media content. Not only do various successful predictors of encoded exposure theoretically reside at different levels of measurement, but it also appears that some of these variables moderate the influence of variables located at a different level. Individual adolescents in the U.S. exert some limited influence over their own exposure to mass media campaigns, but they also appear to be living in a web of influences, ranging from conversations with others to particular features of media content, that affect their memory for campaign material in a variety of ways.

Notes

[1] Eagly and Chaiken (1993) note, for example, that the HSM permits heuristic and systematic processing to occur simultaneously and that heuristic processing, and heuristic cues, can affect systematic processing. Moreover, the HSM holds that motivational variables can not only invite systematic processing but also can affect heuristic processing as well.

[2] See Rowan, Raudenbush, & Kang (1991) for a useful and thorough overview of intraclass correlation and its relevance to multilevel modeling.

[3] The youth and their parents were found by door-to-door screening of a scientifically selected sample of about 34,700 dwelling units for Wave 1 and a sample of 23,000 dwelling units for Wave 2. These dwelling units were spread across about 1,300 neighborhoods in Wave 1 and 800 neighborhoods in Wave 2 in 90 primary sampling units. The sample provided an efficient and nearly unbiased cross-section of America's youth and their parents. Youth living in institutions, group homes, and dormitories were excluded. Among selected youth, the response rate was approximately 91 percent in Wave 1 and 92 percent in Wave 2, meaning that 91 or 92 percent of the youth received parental consent, signed to their own assent, and completed an extended interview.

[4] This random selection was accomplished by first using both dwelling unit identification number and roster identification number as grouping variables (making each respondent into a single group, in other words) and then randomly selecting one case from each created group. SPSS syntax for this operation was adapted from the following SPSS advice web site in February 2002: "http://pages.infinit.net/rlevesqu/Syntax/RandomSampling/Select2CasesFromEachGroup.txt".

[5] The present analyses employ NSPY weights that reflect sample selection probabilities and compensate for non-response (Hornik et al, 2001). As present analyses utilize HLM 5, however, replicate weights adjustment available through other programs, such as WesVarPC, could not be employed. Accordingly, I emphasize those results with $p < .01$, as opposed to results at the conventional .05 level.

[6] For both weekday and weekend watching, the "none" category was assigned "0". For weekday watching, "half-hour or less" was assigned ".5" and, for weekend watching, "less than one hour" also was assigned ".5". For weekday watching, the "about 1 hour" through "about 6 hours" categories were assigned "1" through "6", respectively. The "7 hours or more" category was assigned "8" for weekday watching. For weekend watching, the "1 to 2 hours" through " 9 to 10 hours" categories were assigned "1.5", "3.5", "5.5", "7.5", and "9.5", respectively, and the "11 hours or more" category was assigned "12".

7 The original NSPY questions asked how often the respondent had watched each of the following in the past 30 days: "a music television station, such as MTV, VH1, or TNN (The Nashville Network)", "an all-sports channel, such as ESPN", or "a channel focused on African Americans or Blacks such as BET." Spanish-language interviews also asked how often one had watched "a channel especially for Latinos or Hispanics such as Telemundo, Univision, or Galavision" in the past 30 days. Original response categories included "never", "1 to 4 days", "5 to 14 days" and "15 to 30 days" and were assigned the interval levels of "0", "2.5", "9.5", and "22.5", respectively.

[8] The mean of USEDEPTH was .31, SD = .63, skewness = 1.84. This reflects the fact that most youth report no past marijuana use, though a small number report past regular use. The positive skew suggested the usefulness of a variable transformation. The natural log of USEDEPTH (which we can label LNUSEDEP) demonstrated much less skew than the original variable and was useful for analysis. The skewness of the LNUSEDEP distribution was 1.37 (mean = -8.94, SD = 4.88).

[9] "Never" was recoded into 0, "Once" was recoded into 1, "2 to 3 times" was recoded into 2.5, "4 to 5 times" was recoded into 4.5 times, "6 to 10 times" was recoded into 8, and "More than 10 times" was recoded into 12.

[10] A cut is a transition to a different camera perspective that results in the depiction of a new visual environment or entirely new visual information. The following rules further clarify that notion. Any transition to a new physical environment (one that is not visible in, or contiguous with, the previous shot) counts as a cut. A transition to a close up of a face (at least 1/5th of the screen) also counts as a cut, even if the face was partially visible in the establishing shot and the same environment is depicted. This idea is based on Lang's (personal communication, 2001) recommendation. If transition depicts the exact same room but results in the depiction of an entirely new face in the same room, it will count as a cut the first time that the person (or people) in question appears. Each subsequent repetition of the person will not count as a cut (unless, of course, the environment has changed between shots of the face and the shot with a face now represents a cut from a different visual environment). If the same people are depicted in the exact same room in a sequence of shots that could not physically have occurred without editing, e.g., alternative versions of the same scenario, the first transition to a repeated scenario will count as a cut. Each subsequent repetition in the uninterrupted sequence will not count as a cut. Any transition from whole screen to split screen with different environments depicted is cut. Each new introduction of new scene in each separate screen (in case of split screen) is a cut. Transition to whole screen from split in which one of the scenes is enlarged to become the whole screen is an edit and not a cut. Transition to black (or other color) screen with text is a cut. Transition from one line of text to another, however, is an edit and not a cut. Special effects allow for some transitions in which only a part of a screen display changes, e.g., an abstract image changing one-fourth at a time. In these cases, at least half of the total screen area needs to change to a new image in order to constitute a cut.

[11] Robust standard errors are consistent even when ordinary least squares assumptions about constant variance of outcomes across groups are incorrect.

[12] The degrees of freedom are equal to 18 in this instance because only 19 of the original 23 groups had sufficient data for HLM computation of χ^2 to test random effects. Reported fixed effects and variance components, nonetheless, are based on all data.

[13] One age and one race comparison also suggests significant random effects in table 1.2. None of the content-level variables used for this study, however, produced an alternative model that reduced this additional random coefficient variance for age or race significantly. Future investigation of different content-level variables might account for this coefficient variation.

References

Anderson, J. R. (1983). The Architecture of Cognition. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1990). Cognitive Psychology and its Implications. (3rd ed.). New York: W. H. Freeman.

Bower, G.H. (2000). A brief history of memory research. In E. Tulving & F. I. M. Craik (Eds.), The Oxford Handbook of Memory (pp. 3-32). New York: Oxford University Press.

Brickner, M. A., Harkins, S. G., & Ostrom, T. M. (1986). Effects of personal involvement: Thought provoking implications for social loafing. Journal of Personality and Social Psychology, 51, 763-769.

Broadbent, D. E. (1958). Perception and Communication. New York: Pergamon Press.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical Linear Models: Applications and Data Analysis Methods. Newbury Park, CA: Sage.

Bryk, A. S., & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. American Journal of Education, 97(1), 65-108.

Bryk, A. S., Raudenbush, S. W., Congdon, R., & Seltzer, M. (1986). An Introduction to HLM: Computer Program and User's Guide. Chicago: University of Chicago, Department of Education.

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. Review of Research in Education, 8, 158-233.

Cappella, J. N. (1996). Why biological explanation? Journal of Communication, 46(3), 4-7.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), Unintended Thought (pp. 212-252). New York: Guilford Press.

du Plessis, E. (1994). Recognition versus recall. Journal of Advertising Research, 34(3), 75-91.

Eagly, A. H., & Chaiken, S. (1993). The Psychology of Attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Farris, P. W., & Parry, M. E. (1991). Clarifying some ambiguities regarding GRP and average frequency. Journal of Advertising Research, 31(6), 75-77.

Flin, R. H. (1985). Development of visual memory: An early adolescent regression. Journal of Early Adolescence, 5(2), 259-266.

Hagen, I., & Wasko, J. (2000). Introduction. In I. Hagen & J. Wasko (Eds.), Consuming Audiences? Production and Reception in Media Research (pp. 3-28). Cresskill, NJ: Hampton Press.

Haney, W. (1980). Units and levels of analysis in large-scale evaluation. New Directions for Methodology of Social and Behavioral Sciences, 6, 1-15.

Hornik, R. (1989). The knowledge-behavior gap in public information campaigns: A development communication view. In C. T. Salmon (Ed.), Information Campaigns: Balancing Social Values and Social Change (pp. 113-138). Newbury Park: Sage.

Hornik, R. (1997). Public health education and communication as policy instruments for bringing about changes in behavior. In M. E. Goldberg, M. Fishbein, & S. E. Middlestadt (Eds.), Social Marketing: Theoretical and Practical Perspectives (pp. 45-58). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hornik, R., Maklan, D., Cadell, D., Judkins, D., Sayeed, S., Zador, P., Southwell, B., Appleyard, J., Hennessy, M., Morin, C., & Steele, D. (2000). Evaluation of the National Youth Anti-Drug Media Campaign: Campaign Exposure and Baseline Measurement of Correlates of Illicit Drug Use from November 1999 through May 2000 . Bethesda, MD: National Institute on Drug Abuse.

Hornik, R., Maklan, D., Judkins, D., Cadell, D., Yanovitzky, I., Zador, P., Southwell, B., Mak, K., Das, B., Prado, A., Barmada, C., Jacobsohn, L., Morin, C., Steele, D., Baskin, R., & Zanutto, E. (2001). Evaluation of the National Youth Anti-Drug Media Campaign: Second Semi-Annual Report of Findings . Bethesda, MD: National Institute on Drug Abuse.

Katz, E., & Lazarsfeld, P. F. (1955). Personal Influence. Glencoe, Ill: Free Press.

Lang, A. (1995). Defining audio/video redundancy from a limited-capacity information processing perspective. Communication Research, 22(1), 86-115.

Lang, A. (2000). The limited capacity model of mediated message processing. Journal of Communication, 50(1), 46-70.

Lang, A., Geiger, S., Strickwerda, M., & Sumner, J. (1993). The effects of related and unrelated cuts on television viewers' attention, processing capacity, and memory. Communication Research, 20(1), 4-29.

Leippe, M. R., & Elkin, R. A. (1987). When motives clash: Issue involvement and response involvement as determinants of persuasion. Journal of Personality and Social Psychology, 52, 269-278.

Pan, Z., & McLeod, J. M. (1991). Multilevel analysis in mass communication research. Communication Research, 18(2), 140-173.

Petty, R. E., Cacioppo, J., & Haugtvedt, C. (1992). Involvement and persuasion: An appreciative look at the Sherifs' contribution to the study of self-relevance and attitude change. In D. Granberg & G. Sarup (Eds.), Social Judgment and Intergroup Relations: Essays in Honor of Muzafer Sherif (pp. 147-174). New York: Springer/Verlag.

Petty, R. E., & Cacioppo, J. T. (1986a). Communication and Persuasion: Central and Peripheral Routes to Attitude Change. New York: Springer/Verlag.

Petty, R. E., & Cacioppo, J. T. (1986b). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 19, pp. 123-205). New York: Academic Press.

Petty, R. E., & Priester, J. R. (1994). Mass media attitude change: Implications of the elaboration likelihood model of persuasion. In J. Bryant & D. Zillmann (Eds.), Media Effects: Advances in Theory and Research . Hillsdale, NJ: Lawrence Erlbaum Associates.

Price, V., Ritchie, L. D., & Eulau, H. (1991). Cross-level challenges for communication research. Communication Research, 18(2), 262-271.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. T. (2001). HLM 5 for Windows (Version 5.03). Lincolnwood, IL: Scientific Software International.

Rowan, B., Raudenbush, S. W., & Kang, S. J. (1991). Organizational design in high schools: A multilevel analysis. American Journal of Education, 99(2), 238-266.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. Science, 277(5328), 918-924.

Shoemaker, P. J., Schooler, C., & Danielson, W. A. (1989). Involvement with the media: Recall versus recognition of election information. Communication Research, 16(1), 78-103.

Singh, S. N., Rothschild, M. L., & Churchill, G. A. (1988). Recognition versus recall as measures of television commercial forgetting. Journal of Marketing Research, 25, 72-80.

Stapel, J. (1998). Recall and recognition: A very close relationship. Journal of Advertising Research, 38(4), 41-45.

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2000). Ratings Analysis: The Theory and Practice of Audience Research. Mahwah, NJ: Lawrence Erlbaum Associates.

Wright, C. R. (1986). Mass Communication: A Sociological Perspective. New York: McGraw-Hill.

Zinkhan, G. M., Locander, W. B., & Leigh, J. H. (1986). Dimensional relationships of aided recall and recognition. Journal of Advertising, 15(1), 38-46.

CT&M

Running Head: COEFFICIENT OF CO-TERMINATION

Assessing Co-termination Between Coders in Unitizing Textual Data: A Multi-response Randomized Blocks Permutation Approach

Li Cai

School of Journalism and Communication

The Ohio State University

Student Paper

Li Cai recently graduated from the master's program in Journalism and Communication at the Ohio State University, and will soon join the Ph.D. program in Quantitative Psychology at the University of North Carolina – Chapel Hill. Address: L. L. Thurstone Psychometric Lab., Department of Psychology, UNC – Chapel Hill, CB #3270 Davie Hall, Chapel Hill, NC 27599-3270. Phone: 614-596-9197. Email: lcai@email.unc.edu.

Abstract

The assessment of intercoder agreement in the unitizing phase of content analyses has long been overlooked. In particular, little attention has been paid to the issue of co-termination. Although multiple-coder kappa can be used for the purpose of summarizing the agreement of co-termination, its conservativeness often results in gross underestimates. A new family of coefficients based on Multi-response Randomized Blocks Permutation procedure is presented here and numerical examples are given.

Introduction

Content analysis is a quantitative research methodology widely employed in the field of communication. Berelson's (1952) often cited definition of content analysis as an objective, systematic, and quantitative endeavor to describe the content of communication messages clearly endorsed the usefulness of content analysis to communication scholars. In a recent "content analysis of content analyses" of articles published in *Journalism and Mass Communication Quarterly* between 1971 and 1995, Riffe and Freitag (1998) located 486 full-length reports using content analysis, comprising of roughly one fourth of the total number of articles published. An earlier study by Wilhoit (1984) suggested that more than 20 per cent of theses and dissertations listed in *Journalism Abstracts* used content analysis. Moffett and Dominick (1987) reported a similar result that 21 per cent of the articles published in *Journal of Broadcasting* between 1970 and 1985 employed content analysis.

As Krippendorff (1980) succinctly remarked, making "inferences from essentially verbal, symbolic, or communicative data" has always been at the heart of content analysis (p. 20). In order for scientific inferences to be valid, one must first ascertain the reliability of the research instrument. Just as chemists could ill-afford an uncalibrated balance in a chemical experiment, one could hardly imagine living a life in the communication scholarship without assessing the reliability of content analysis. Of course, if a stream of messages is to be analyzed by a well-designed computer program, one could probably worry less about reliability, but in most instances, if not all, content analyses still require much human labor, and thus the probable errors of human analysts become almost inevitable.

The standard process of content analysis, as described in most introductory communication research methods textbooks (e.g. Frey, Botan, & Kreps, 2000; Stewart, 2002; Wimmer & Dominick, 1994) essentially involves a *coding* process. Guetzkow (1950) observed that any transformation of qualitative data into a form "susceptible to quantitative treatment constitutes *coding*" (p. 47). He further emphasized that the coding process could be broken down into two related phases, that of separating the qualitative material into units, and that of classifying the unitized data into established categories. The former is often termed *unitizing*, and the latter *categorizing*. The two processes are integral elements of content analysis yet require different strategies of reliability assessment. One term that needs a little clarification is *categorizing*. In many practical situations the coded units are indeed later classified into categorical sets, but this is not necessarily true. Coded units may be rated on an ordinal/interval or ratio scale in subsequent analyses. The term *categorizing* will remain in use in the paragraphs to follow, but without any implication of merely categorizing the coded units into qualitative (nominal) sets.

In the coding process, usually a set of human *coders* or *judges* are involved. The assessment of reliability of the content analysis thus becomes an assessment of the reliability of the coders, even though this is not a sufficient condition for the entire content analysis study to be reliable, the coding process is of such importance that low intercoder reliability would render all subsequent analyses meaningless, because low intercoder reliability would suggest that the obtained results were largely not replicable (Krippendorff, 1980, p. 131). Ideally, the coders should be trained to rate or judge the content independently and yet to arrive at the same ratings in precisely in the same manner as intended by the coding scheme. Intercoder reliability is established when the same pieces (possibly a very large number) of content yield same ratings from independent coders using a common data language (Krippendorff, 1980, p. 133). Formally, the term intercoder reliability should be more appropriately termed *intercoder agreement* (see Lombard, Snyder-Duch, & Bracken, 2002), but the two terms will nevertheless be used interchangeably in the pages to follow since given the present context the means of two terms are not much different.

Another distinction that should be made is the "depth" at which the messages are to be coded. Berelson (1952) clearly intended content analysts to deal only with the manifest content, i.e. the information "as is," without invoking additional mental efforts of the coders to discover the latent content or the implied meaning. However, unless the research question can be easily answered by simply counting the number of words in a newspaper article or the number of occurrences of the names of candidates in pre-election news coverage – which can be quite easily done with a computer – the coding process will often require the coders to make subjective judgments. Under those circumstances, readers of the research would demand the researchers to show that "those judgments, while subjectively derived, are shared across coders," which again confirmed the practical necessity of establishing intercoder agreement in content analysis (Potter & Levine-Donnerstein, 1999, p. 266).

Having illustrated the importance of intercoder agreement, the current status of correctly using and reporting intercoder agreement measures in communication journals is quite alarming. Riffe and Freitag (1997) found that only half of the 486 articles published in *Journalism and Mass Communication Quarterly* between 1971-1995 reported intercoder reliability. A recent study by Lombard et al. (2002) searched virtually all content analysis articles indexed in *Communication Abstracts* from 1994 to 1998. Of the 200 articles they found, only 69% ever mentioned intercoder reliability, and usually the methods for computing intercoder reliability were not reported. Of the 44% of all articles that did report the names of the specific methods, more than half of them relied on liberal indices that are not chance-corrected, such as percent agreement, which seriously undermined the effort of computing and reporting reliability coefficients (Lombard, et al., 2002, p. 596).

Given the current undesirable state of affairs of appropriately using and reporting intercoder agreement indices in the communication scholarship, the next section shall explicate intercoder agreement in the context of a two-stage coding process, namely, unitizing and categorizing. The importance of *co-termination* when unitizing textual data shall be presented.

Co-termination

When Guetzkow (1950) wrote about unitizing and categorizing, he presented a convincing case that in order for the entire content analysis to be reliable, one has to regard the assessment of the overall intercoder agreement as a two-stage process. Ideally one should compute agreement measures for unitizing first and then calculate the agreement indices for categorizing, with the "overall" reliability referring to the combined intercoder agreement in both stages. One should note that this overall reliability does not always have to be expressed in quantitative terms. It is possible that a particular content analysis consists merely of categorizing existing units, and then this two-stage notion would not be relevant. However, there are times when unitizing is a must, and under such circumstances, the intercoder agreement of unitizing becomes crucial. This paper does not attempt to develop any new agreement indices for the categorizing phase, as there are established methods already. Instead, the aim is on how the agreement of unitizing can be better summarized, and this goal cannot be achieved without first understanding the complexities of intercoder agreement in the unitizing phase.

The problem of agreement of unitizing focuses on how independent coders choose breaking points at various places in a continuous *segment* of textual content, be it a sentence, a paragraph, an article, or an entire television show. The segments are assumed to be clearly delineated from one another and are usually naturally given. This assumption is not unfair because most of the qualitative content that can serve as segments for coding has unambiguous endpoints. For example, if a newspaper article is chosen as a segment, where it ends is crystal clear. It is further assumed that segments are of two types, discrete and continuous. The reader should take note that such a distinction is probably quite artificial. The only reason behind such a distinction is the mathematics.

Discrete segments are composed of a finite number of *elements*. Defining what is an *element* is difficult, because it ultimately depends on the research question, and how detailed the researcher would like the content analysis to be, but an example should help illustrating this concept. For instance, a sentence from an online chat transcript is selected: "Apparently, from what I read, they haven't identified the dead body yet." It is convenient to define a word – anything in between two spaces – as an *element*. Therefore, this is a segment containing 12 elements. Thus defined, unitizing becomes an operation of grouping elements into units, and intercoder disagreement arises when coders define the groups differently.

The idea of elements is not applicable to continuous segments. For instance, a researcher may want to unitize audio/video recordings. It is probably hard to define what an element is within a continuous stream of audio/ video recording, but it is easy to deal with the relative *length* of a unit, perhaps expressed in terms of time. One can imagine the coder using a stopwatch to record the lengths of units, and intercoder disagreement occurs when the coders come up with different length readings. The idea of length is widely applicable and it is easy to see that one can actually express the discrete type of unitizing using lengths as well.

Consider this example: ABCD, a discrete segment of 4 elements, was to be coded by two judges by putting slashes at the breaking points. Judge 1 gave: A/B/CD, and judge 2 gave: A/BC/D. They both came up with three units for this segment, and they were said to be co-terminus for the first unit. The reliability data, using discrete terms, can be thought of as a set of binary streams: 1 1 0, for judge 1; and 1 0 1 for judge 2. The number of entries in the binary stream is the number of elements minus 1, representing the number of possible breaking points.

The 1s in a stream signify observed breaking points. The same data can be expressed in terms of lengths: 1 1 2 for judge 1; and 1 2 1 for judge 2. The numbers correspond to the number of elements in a particular unit, and the total number of entries equals the number of units.

Having defined the terms, it is natural to introduce the concept of *co-termination* and review what Guetzkow (1950) recognized as the two kinds of errors likely to be present when unitizing a stream of content: (1) failure to agree on the breaking points between the units, and (2) failure to attain the same number of units (p. 54). Co-termination, or co-terminability, a term introduced but not clearly defined in Guetzekow (1950), refers to the agreement among pairs of coders to break a given segment of content at the same points into the same number of smaller units. Note that this definition essentially contains two components: (1) the agreement on the breaking points, and (2) the agreement on the number of units. Such a definition of co-termination is said to be in a "strong" form because there will be perfect agreement of unitizing among coders when the strong form of co-termination is achieved. It is the necessary and sufficient condition for a "weaker" form to exist because it is possible that a pair of coders agree partially, such as for the first unit in the afore mentioned example, on how to choose the breaking points and yet at the same time do not agree on how many units there are in the segment of content. An example should help illustrating this point. Suppose two coders were instructed to break an article into smaller units containing one or more paragraphs. The two coders started out in perfect agreement as to how to group the paragraphs into units up to a certain paragraph after which things started to fall apart. As a result, the numbers of units were different, and certainly by definition of strong co-termination, they failed to achieve agreement. However, one has to acknowledge that at least the two agreed somewhat in the beginning, and a good agreement measure should give partial credit to what they agreed upon. It is conceivable that any measure of agreement based on the strong form of co-termination would necessarily be a conservative one and thus the existence of a weak form of co-termination is not an idea plucked out from the thin air.

The weak form of co-termination essentially depends on the sequential nature of content streams, i.e. one can only start unitizing from the beginning of a segment and proceed as the stream goes. Of course, going backwards from the end is not impossible, but this is makes little difference because one can then define the end as the beginning. Expressed in discrete terms, the weak form of co-termination between two-coders is defined as choosing breaking points so that at least the two coders grouped one set of elements in the same manner. Consider the first example again: a segment – ABCD, with 4 elements, and 3 coders were to unitize it. The result happened to be as follows: coder 1 – A/B/CD, coder 2 – AB/C/D, and coder 3 – A/B/C/D. There are three distinct pairs of coders: 1 vs. 2, 2 vs. 3, and 1 vs. 3. Clearly, none of the pairs achieved co-termination if the strong definition is used. Coders 1 and 2 gave the same number of units but were not co-terminus. Although coders 1 and 3 gave different numbers of units (3 and 4, respectively), they actually attained the weak form of co-termination for the groupings of elements A and B into the first and second unit. For coders 2 and 3, they achieved co-termination for C and D. The basic conceptualization of the measurement of co-termination would be to employ the strong definition when the coders agree on the number of units and to use the weak form when the numbers of units are different.

It is worthy of pointing out that according to Hubert (1977) there are three definitions of agreement when the number of coders goes beyond two: DeMoivre's definition, target-rater definition, and pair-wise definition. The first one refers to the unanimous agreement of all coders, and the second one refers to the joint agreement of all other coders with a "target-rater" who provides the "true" rating, and the third, which is also what is implied in the definition of co-termination, refers to the agreement between any pairings of coders. It is easy to see that

DeMoivre's definition tends to yield the most conservativeness. Most of the popular intercoder agreement indices that can handle three or more coders use the pair-wise definition, as does the new coefficient to be proposed in subsequent sections.

Having defined what co-termination is, it is not difficult to infer that the mere agreement on number of units does not imply co-termination. As to the relative importance of the two, Guetzkow (1950) remarked that the failure to achieve "co-terminability" is less likely to lead to confusions and low intercoder reliability in the subsequent categorizing of the coded units (p. 55). There is absolutely nothing wrong with this argument, because how far reliability assessment should go is a practical matter related to the nature of the specific study at hand. If the unit boundaries are relatively clear, or if slight inconsistencies in co-termination do not significantly affect the subsequent use of the coded units, one could worry less about co-termination and focus more on achieving a high level of agreement on the number of units. However, there are certain times when disagreement in co-termination may lead to different interpretations of the same data, even though the number of units are the same across coders. For instance, if two coders were to divide the sentence "Apparently, from what I read, they haven't identified the dead body yet," and the coders agreed that it contained two units, but the first coder put the division mark right after "apparently," while the second put it after "read." The interpretation of the two units would necessarily be different, because a stand-alone "apparently" would suggest confirmation, while "apparently, from what I read" would refer to the clear inferences that the chat user could make from what he or she read. This example is only a very trivial one. What is important is to realize that the mere agreement on number of units does not automatically imply reliability of unitizing.

Still using the previous example, suppose that the first coder divided the sentence after both "apparently" and "read," and the second coder only divided the sentence after "apparently," the number of units for the two coders are 3 and 2, respectively, and there seems to be much disagreement between the two, but in fact they did achieve co-termination, at least for the first unit. Given such results, at least the interpretation for the first unit – "apparently," is unambiguous. In the next section, the five most widely used indices of intercoder agreement shall be briefly reviewed. Most of them are intended for bivariate nominal level coding, and for discrete reliability data (binary streams) between two coders Cohen's κ can be used, but only to a limited extent because of the resulting gross underestimate of reliability, which will become clear in the sections to follow. However, for continuous content, no current indices are directly applicable and a new generalized measure based on Multi-response Randomized Blocks Permutation procedures (Mielke & Iyer, 1982) shall be presented.

Popular Indices of Intercoder Agreement

Percent Agreement and Holsti's Method

This is perhaps the most easily understood method for calculating intercoder agreement. It is simply the "percentage of all coding decisions made by pairs of coders on which the coders agree" (Lombard, et al., 2002, p. 590). This is not a chance corrected measure, and Krippendorff (1980) illustrated how chance could artificially inflate percent agreement with a neat example (pp. 133-135). In general, using percent agreement is a very poor practice that inflates reliability, and is not applicable to other higher levels of measurement than nominal level coding.

Holsti (1969) proposed a variation of the percent agreement measure, which is the same as percent agreement when two coders are coding the same segments of content. This is still not a chance-corrected measure and it suffers from the same drawbacks as percent agreement. It is interesting to note that even though some statisticians have argued against the use of chance-corrected measures (e.g., Goodman & Kruskal, 1956), supporters of chance-corrected measures "far outweigh detractors" (Berry & Mielke, 1988, p. 922).

Scott's π

This is a chance-corrected index first introduced by Scott (1955) primarily in the context of coding qualitative data obtained from surveys. In its original form, this index is only applicable to univariate nominal level coding and accommodates only two coders, although it is worth mentioning that Craig (1981) has given an extension of Scott's π to the case of multiple coders. Scott'*s* π is the first coefficient that considers both the number of categories and the marginal distributions, i.e. how the two coders distribute their classifications of the units. However, the π coefficient not only assumes that the column and row marginal distributions are identical to the "true" proportions, but also takes it a step further by assuming that the two coders share the same marginal distributions. Given the context of survey research, the former assumption is not unreasonable, as the "true" proportions are usually obtainable, and this assumption has given Scott's π a distinct edge over similar coefficients like Cohen'*s* κ, because π can still be computed when the two coders have coded different subsets of the content, while computation of κ requires that the pair of coders have coded the same units (Craig, 1981, p. 261). However, it is precisely the latter assumption of π that is more problematic. As Cohen (1960) pointed out, "one source of disagreement between a pair of judges is precisely their proclivity to distribution their judgments differently over the categories" (p. 41). Furthermore, the "true" proportions are not always available, thus making such an unrealistic assumption only hinders the general practicability of the π coefficient.

Cohen's κ

Cohen's (1960) κ is defined in much the same way as Scott's π. Usually it is assumed that two coders independently classify each of the n units into one of c established categories. The layout for computing such bivariate nominal level intercoder agreement as κ essentially involves the construction of a two-way cross-classification table, with entries in the table being the proportion of observations falling into one of the c by c cross-classifications. The marginal distributions are simply the column and row sums. The κ coefficient, and its variants for bivariate nominal data usually assumes the form of a ratio between observed and expected proportions $\kappa = (P_o - P_e)/(1 - P_e)$, with P_o given by the sum of the diagonal elements of the c by c cross-classification table, and P_e is found by first multiplying each column marginal with its associated row marginal and then taking the sum of the products.

Cohen's κ has enjoyed continued development by psychological methodologists. Cohen (1968) himself introduced a weighting procedure that accounts for the differential severity of disagreements. Fleiss (1971) gave its extensions to the case of multiple raters. Fleiss and Cohen (1973) established the equivalence of weighted κ and the intra-class correlation coefficient. Hubert (1977) introduced the underlying mathematical model of matching distributions in probability theory to users of the κ coefficient. Fleiss, Nee, and Landis (1979) worked out κ's asymptotic variance. Conger (1985) extended it to measure agreement over time for continuous scales. As mentioned afore, κ can be used to assess co-termination for two coders and discrete type of unitizing, but results in an underestimate.

Krippendorff's α

When the number of coders is exactly two with nominal level coding assumed, Krippendorff's (1970) α coefficient is identical to Scott's π (cf. Krippendorff, 1980, p. 138). What makes the α coefficient more appealing than its competitors is that it offers an easy extension to measure the agreement of higher levels of measurement and of multiple coders. Recall that Guetzkow (1950) described the two kinds of errors in unitizing textual data. It appears that Krippendorff's α coefficient may well serve the purpose of calculating the intercoder agreement of the number of units of a given segment of content, but it is still lacking in its ability to detect co-termination for the continuous case.

Multi-response Randomized Blocks Layout and Intercoder Agreement

This section describes some of the details of the Multi-response Randomized Blocks Permutation procedure (MRBP) relevant to the assessment of agreement, and the details of the equations can be skipped without loss of continuity.

Multi-Response Permutation Procedure (MRPP) is a versatile analytic framework first outlined in Mielke, Berry, and Johnson (1976) as a robust and powerful tool for analyzing multivariate data from randomized experiments using the average within-group distance as the test statistic (descriptions of MRPP are given in Appendix I). The usual form of the symmetric distance function between two multi-dimensional responses is given by

$$\Delta_{I,J} = \left[\sum_{c=1}^{r} (x_{cI} - x_{cJ})^2 \right]^{v/2}, \qquad (1)$$

where ($x_{1I}, \ldots, x_{rI}$) denote one r-dimensional response, I and J are distinct integers from 1 to N, and N is the total number of responses, i.e. the total sample size. It is easy to see that the distance between two multivariate responses is a power function of the summation of the squared distances between each dimension and therefore the choice of v gives rise to a variety of distance functions. The value of v determines the analysis space of the test and choice is somewhat arbitrary, but the most widely used two are $v = 1$ and $v = 2$, which corresponds to metric Euclidean distance and non-metric squared Euclidean distance. Some of the most widely employed tests such as the t-test, *ANOVA*, and their multivariate counterparts – Hotelling T^2, and Bartlett-Nanda-Pillai trace test in *MANOVA* all use squared Euclidean analysis space. Berry and Mielke (1988) pointed out that the choice of squaring the distances is "questionable at the best" (p. 922). They suggested $v = 1$ be used at all times based on its robustness against outliers, but

Janson and Olsson's (2001) modified statistic uses $v = 2$ and their main argument for the more conventional metric is the ease of interpretation. For binary streams of discrete reliability data, the choice does not matter because the square root of 1 is still 1, but for continuous content, as the reader will see later, $v = 2$ is the sometimes the only choice because of the vast reduction in computation time.

Multi-Response Randomized Blocks Permutation (MRBP) is a variation of MRPP statistic when a blocking variable is added into a one-way design. It is first introduced by Mielke & Iyer (1982) as a supplement to MRPP. In its original formulation, MRBP defines a b-block by g-treatment randomized experiment and within each block there is only one r-dimensional observation per treatment, taken as $n = 1$ for each cell. Using the MRBP layout, Berry and Mielke (1988) provided a re-formulation of Cohen's κ and a natural extension of κ to multiple coders as well as to higher levels of measurement.

In brevity, the original cross-classification layout of κ is transformed into a b-block by g-treatment MRBP layout. Assuming that two observers independently coded each of the g units into one of r categories. The usual cross-classification layout of κ would be an r by r table with the entries in the table being the proportions of cross-classifications in particular cells. The MRBP layout, however, would be a 2-block by g-treatment design with r-dimensional responses and v set to 1 when calculating distances. The extended measure of agreement is given by the equation

$$\kappa = 1 - \delta / \mu_\delta, \tag{2}$$

where δ denotes observed disagreement and μ_δ denotes expected proportion of disagreement by chance. Because MRBP is a based on permutation, μ_δ is found by permuting the data within each block across treatments. Such a formulation makes the extension of κ to higher levels of measurement and multiple coders easy.

Generally, assuming b coders independently rate g units of content, let $[x_{ijk}]$ denote the elements in an r-dimensional response vector (when $r = 1$, this is a scalar) from coder i for unit j, where $i = 1, \ldots, b, j = 1, \ldots, g$, and $k = 1, \ldots, r$, the within-treatment distance function is given by

$$\Delta_{p,q} = \left[\sum_{c=1}^{r} (x_{pjc} - x_{qjc})^2 \right]^{v/2}, \tag{3}$$

and the average within-treatment distance for all distinct pairs of coders is given by

$$\delta_{obs} = \left[g \binom{b}{2} \right]^{-1} \sum_{j=1}^{g} \sum_{p<q} \Delta_{p,q}, \tag{4}$$

where $p < q$ denotes the sum over all p and q such that $1 \leq p < q \leq b$, and basically this is to ensure that a response vector is not compared with itself. The definition of μ_δ reflects the addition of blocks because unlike MRPP, in randomized blocks designs data cannot be permuted across the blocks. Therefore the maximum number of permutations is $M = (g!)^b$ – the total number of permutations within each block to the bth power. Assuming the M permutations are equally probable, a theoretical definition of chance disagreement is given by

$$\mu_\delta = M^{-1} \sum_{i=1}^{M} \delta_i, \tag{5}$$

However, one does not need to enumerate all M permutations to arrive at μ_δ, a more efficient working formula for μ_δ is available due to the fact that the first moment of the permutation distribution is a constant multiple of g^2 elementary calculations (see Mielke & Iyer, 1982).

Using similar notations as in equations (3) and (4), then the within-and-between-units distance function between two rating vectors is given by

$$\Delta_{pi,qj} = \left[\sum_{c=1}^{r} (x_{pic} - x_{qjc})^2 \right]^{1/2}, \tag{6}$$

and the following equation gives the chance disagreement

$$\mu_g = \left[g^2 \binom{b}{2} \right]^{-1} \sum_{i=1}^{g} \sum_{j=1}^{g} \sum_{p<q} \Delta_{pi,qj} . \tag{7}$$

Equation (7) seems to be quite complicated. However, it is nothing but the average distance between any distinct pairings of response vectors. Berry and Mielke (1988) have named their extended κ coefficient as $\mathscr{R}$, and have established the equivalence of this statistic with other known measures.

Formulation of the Proposed Coefficients of Co-termination

Assuming two coders are present, and they have broken a 7-word sentence into 3 units. The analysis of intercoder agreement, using discrete terms, may be expressed as a 2 block by 6 treatment MRBP layout. The entries are just 0s and 1s, and the design is summarized in Table 1.

Table 1

Example dataset

	Treatments					
Blocks (Coders)	1	2	3	4	5	6
1	1	0	0	0	1	0
2	1	0	0	0	0	1

If this sentence is of the form ABCDEFG, then the codings in Table 1 are: A/BCDE/FG for coder 1, and A/BCDEF/G for coder 2. If the cross-classification layout of κ is used, the design should be a 2 by 2 cross-classification table, and it would look like Table 2.

Table 2

Cross-classification Layout

	Coder 1		
Coder 2	1	0	Sums
1	1	1	2
0	1	3	4
Sums	2	4	6

Cohen's κ can be calculated from Table 2 using the usual way.

$$\kappa = \frac{P_o - P_e}{1 - P_e} = \frac{(1/6 + 3/6) - [(2/6 \times 2/6) + (4/6 \times 4/6)]}{1 - [(2/6 \times 2/6) + (4/6 \times 4/6)]} = .25$$

More generally, let $[x_{ij}]$ represent the value (0 or 1) from the cell corresponding to the *i*th block and *j*th treatment, where $i = 1, \ldots, b$, and $j = 1, \ldots, g$, the symmetrical MRBP distance function between any two cells $[x_{ij}]$ and $[x_{pq}]$ in a table similar to Table 1 can be simplified to

$$\Delta_{ij,pq} = (x_{ij} - x_{pq})^2. \tag{8}$$

Using equations (3) – (7), a reformulated κ can be expressed as

$$\kappa' = 1 - \frac{\delta'}{\mu_\delta'} \tag{9}$$

where δ' is the average within-treatment disagreement and μ_δ' – the expected disagreement – can be found by averaging over all δ's obtained from permuting data within blocks. Table 3 is an example of a possible permutation. For instance, in Block 1, the 1s originally in the 1st and 5th treatments are swapped into the 2nd and 3rd places. For this permutation $\delta' = 4/6 = 2/3$.

Table 3

A Possible Permutation

	Treatments					
Blocks (Coders)	1	2	3	4	5	6
1	0	1	1	0	0	0
2	1	0	0	0	0	1

To summarize, δ' can be calculated as (0+0+0+0+1+1)/6 = .333, μ_δ' = .444, and κ' is 1 – .333/.444 = .25, which is exactly the same as using the cross-classification table, but the MRBP approach can be easily extended to multiple coders.

The problem with this approach, as the reader probably has already noticed, is an underestimate of reliability. Without calculating any statistics, a visual examination of the codings: A/BCDE/FG for coder 1, and A/BCDEF/G for coder 2, reveal the fact that the codings are not much different yet the agreement measure indicates that it is only 25% agreement above chance, a value too low by any standards. Therefore, a remedy shall be presented and it makes use of the notion of continuous content.

One can express the data in Table 1 using lengths and the result is summarized below. For the moment, the reader is asked to ignore the lines corresponding to the "Cumulative" lengths. The usefulness of these identities will become clear later.

Table 4

Continuous measurement

	Treatments		
Blocks (Coders)	1	2	3
1	1	4	2
Cumulative 1	1	5	7
2	1	5	1
Cumulative 2	1	6	7

One can essentially apply equations (3) – (7) on the two blocks and follow computational formulae in Mielke and Iyer (1982) to obtain the reliability coefficient. With $v = 1$, $\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - .667 / 1.778 = .625$. With $v = 2$, $\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - .666 / 5.111 = .87$. One can see that by using continuous content, the agreement index is increased quite a bit. However, the direct application of continuous content is quite problematic given that coders usually do not agree on the number of units either. Not only the computational formulae are rendered useless, there are conceptual problems too. Consider the following coding:

Table 5

Continuous measurement

	Treatments						
Blocks (Coders)	1	2	3	4	5	6	7
1	10.3	19.7	10.1	9.8	4.6	5.0	10.5
Cumulative 1	10.3	30.0	40.1	49.9	54.5	59.5	70.0
2	10.2	19.8	8.5	12.0	9.5	10.0	–
Cumulative 2	10.2	30.0	38.5	50.5	60.0	70.0	70.0

The first coder came up with 7 units for this continuous piece of content and the second one came up with only 6. One can easily think of replacing the missing cell in the last column with zero and applying the computational formulae thereafter, but then a theoretical problem arises because when the permutation within the second block is conducted, that imputed zero may appear, for example, in the 3rd column. It makes little sense because one can hardly imagine a unit of length zero in between two other units of positive lengths. After all, the communication content is sequential stream that does not stop until the endpoint.

This problem can be offset by fixing the trailing zero(s), should there be one or more missing cells in the last a few columns, when conducting the within block permutations. Therefore, the total number of possible permutations in the given example is only (7!)(6!) = 3,628,800, instead of $(7!)^2 = 25{,}401{,}600$, as the trailing zero at $[x_{27}]$ will remain un-permuted. More generally, the total number of permutations is given by

$$M = \prod_{j=1}^{b} g_j!, \tag{10}$$

and when all g_j's are equal, equation (10) is equal to $(g_{max}!)^b$. This change essentially reflects the usefulness of the so-called *reference subsets* described in Edgington (1987). If the set of $(g_{max}!)^b$

data permutations is taken as the primary reference set, then the computation of μ_δ under the condition when all g_js are equal would be using the reference distribution for the general-null hypothesis, whereas when not all g_js are equal, and thus equation (10) yields a smaller value than $(g_{max}!)^b$, the computation of μ_δ would be comparable to the test of a restricted mull hypothesis (see Edgington, 1987, pp. 305-316).

When not all g_js are equal, it is useful to define the following computational expressions for computer implementation. A GAUSS (Aptech Systems Inc., 1997) procedure which implements these formulae is in Appendix II.

Let x_{pi} denote a row of g_i unit length data from coder i, where $p = 1, \ldots g_i$, and let M be given as in equation (10), define:

$$C_1(i,j) = \frac{M}{\max(g_i, g_j)}, \qquad (11)$$

$$C_2(i,j) = M\left[1 - \frac{\min(g_i, g_j)}{\max(g_i, g_j)}\right], \qquad (12)$$

$$C_3(i,j) = \begin{cases} i & \text{if } g_i > g_j, \\ j & \text{if } g_i < g_j. \end{cases} \qquad (13)$$

$$D_1(i,j) = \sum_{p=1}^{g_i} \sum_{q=1}^{g_j} \Delta_{pi,qj}, \qquad (14)$$

$$D_2(i,j) = \sum_{p=1}^{g_{C_3(i,j)}} x_{pC_3(i,j)}, \qquad (15)$$

$$E(i,j) = C_1(i,j)D_1(i,j) + C_2(i,j)D_2(i,j), \qquad (16)$$

$$\mu_\delta = M^{-1} \sum_{i<j} E(i,j), \qquad (17)$$

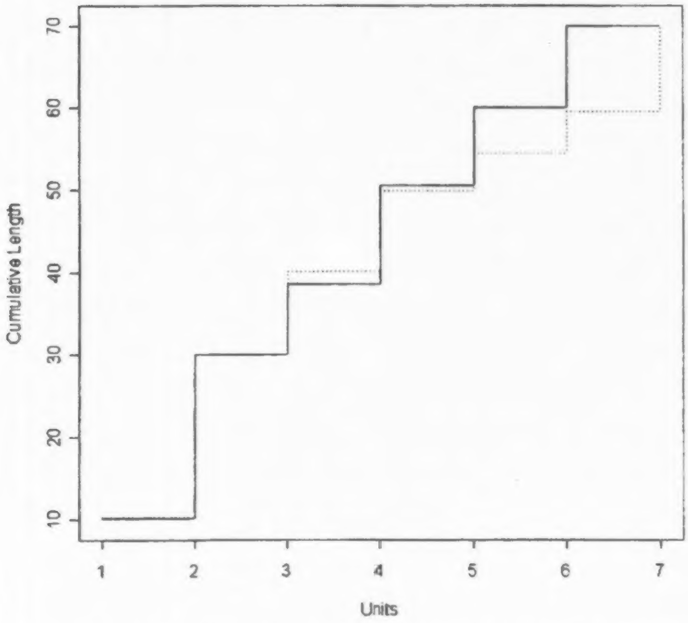
where $i < j$ denotes the summation over all i and j such that $1 \le i < j \le b$.

With $\nu = 1$, $\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - 3.486/ 5.265 = .338$. With $\nu = 2$, $\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - 23.811/ 50.174 = .525$.

A conceptually much simpler way makes use of the "Cumulative" lengths. Borrowing the concept of empirical cumulative distribution function (ECDF) in elementary mathematical statistics, one can envision the disagreement between two coders when unitizing continuous content as the difference between two cumulative length functions. A plot should help illustrating this point.

Figure 1

Two Cumulative Length Functions



The dotted line corresponds to coder 1 and the solid line corresponds to coder 2. In Table 5, when coder 2 has used up all available content, the cumulative length is 70.0 and it remains at 70.0 regardless of how many missing cells there may be.

Using similar notations as above, let $[x_{ij}]$ represent the value from the cell corresponding to the *i*th Block and *j*th Treatment, where $i = 1, \ldots, b$, and $j = 1, \ldots, g$, where g is the maximum number of units given by the coders. One may consider $g = \max(g_1, g_2, \ldots, g_b)$, and the squared distance ($v = 2$) between two cumulative lengths corresponding to cells $[x_{ij}]$ and $[x_{kj}]$ in a table similar to Table 5 can be expressed as

$$\Delta_{ij,kj} = \left[\left(\sum_{p=1}^{j} x_{ip}\right) - \left(\sum_{p=1}^{j} x_{kp}\right)\right]^2 = \left[\sum_{p=1}^{j} (x_{ip} - x_{kp})\right]^2, \qquad (18)$$

and the average observed disagreement is

$$\delta = \binom{b}{2}^{-1} \sum_{i<k} \sum_{j=1}^{g} \Delta_{ij,kj}, \qquad (19)$$

where $i < k$ denotes the summation over all i and k such that $1 \le i < k \le b$.

The computational formulae for expected disagreement are so cumbersome that precludes presentation here due to the complexities involved. However, a GAUSS (Aptech Systems Inc., 1997) procedure that implements this method is given in Appendix III. If one is willing to assume equal probability being placed on every permutation of these cumulative lengths, everything else then follows as what Berry and Mielke (1988) described.

The numerical results for the example datasets in Tables 4 and 5 are as follows:

$\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - 1 / 10.222 = .902.$

$\kappa = 1 - \delta_{obs} / \mu_\delta = 1 - 143.43 / 551.197 = .74.$

Tests of Significance

Since κ is merely a linear function of δ_{obs}, a test of significance of κ is equivalent to the test of δ_{obs}. Mielke and Iyer (1982) gave formulae for the first three moments of the MRBP null distribution, and using the mean and variance, δ_{obs} can be standardized and the associated probability of δ_{obs} can be approximated via a Pearson type III distribution (see Mielke and Berry, 2001). This *p*-value is associated with the test whether κ is significantly different from zero. Since no random sampling assumption is involved, this test of significance is non-asymptotic and is different from what most computer packages do. Each one of the *n* segments in a reliability study would therefore have a *p*-value, and by looking at the set of *p*-values, the researcher should be able to infer whether the coders' overall agreement is due to chance or not.

A test of significance may also be conducted for the coefficient that uses the cumulative lengths via a random sample of all possible permutations (see Edgington, 1987). The exact moments of the null distribution can also be derived along the same lines as equations (11) – (17), but would be very cumbersome.

In summary, the assessment of co-termination is an important issue for content analysts. Coefficients of co-termination under various circumstances were considered, including discrete and continuous content, binary stream data and length data, $v = 1$ and $v = 2$, and the special case involving cumulative lengths. Generally speaking, for discrete data, multiple-coder κ can be used directly, but results in underestimate of agreement. Choosing $v = 2$ over $v = 1$ increases agreement. The coefficients of the continuous type can handle unequal number of units, especially the coefficient that makes use of cumulative lengths. It is also the least conservative among all indices. Depending on the nature of the study, researchers now possess a family of intercoder agreement indices for unitizing textual data, based on the Multi-response Randomized Blocks Permutation procedure.

References

Aptech Systems, Inc. (1997). GAUSS (version 3.2.30) [computer program]. Maple Valley, WA: Author.

Berelson, B. (1952). *Content analysis in communication research*. Glencoe, IL: Free Press.

Berry, K. J., & Mielke, P. W. (1983a). Computation of finite population parameters and approximate probability values for multi-response permutation procedures (MRPP). *Communications in Statistics – Simulation and Computation, 12*, 83-107.

Berry, K. J., & Mielke, P. W. (1983b). Moment approximations as an alternative to the *F* test in analysis of variance. *British Journal of Mathematical and Statistical Psychology, 36*, 202-206.

Berry, K. J., & Mielke, P. W. (1988). A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement, 48*, 921-933.

Berry, K. J., & Mielke, P. W. (1992). A family of multivariate measures of association for nominal independent variables. *Educational and Psychological Measurement, 52*, 41-55.

Berry, K. J., & Mielke, P. W. (1997a). Agreement measure comparisons between two independent sets of raters. *Educational and Psychological Measurement, 57*, 360-364.

Berry, K. J., & Mielke, P. W. (1997b). Measuring the joint agreement between multiple raters and a standard. *Educational and Psychological Measurement, 57*, 527-530.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*, 213-220.

Conger, A. J. (1985). Kappa reliabilities for continuous behaviors and events. *Educational and Psychological Measurement, 45*, 861-868.

Craig, R. T. (1981). Generalization of Scott's index of intercoder agreement. *Public Opinion Quarterly, 45*, 260-264.

Edgington, E. S. (1987). *Randomization tests* (2nd ed.). New York, NY: Marcel Dekker.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*, 378-382.

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement, 33*, 613-619.

Fleiss, J. L., Nee, J. C. M., & Landis, J. R. (1979). Large sample variance of kappa in the case of different sets of raters. *Psychological Bulletin, 86*, 974-977.

Fowler, G. L. (1986). Content and teacher characteristics for master's level research course. *Journalism Quarterly, 63*, 594-599.

Frey, L. R., Botan, C. H., & Kreps, G. L. (2000). *Investigating communication: An introduction to research methods* (2nd ed.). Boston: Allyn & Bacon.

Goodman, L. A., & Kruskal, W. H. (1956). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*, 732-764.

Guetzkow, H. (1950). Unitizing and categorizing problems in coding qualitative data. *Journal of Clinical Psychology, 6*, 47-58.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

Hubert, L. (1977). Kappa revisited. *Psychological Bulletin, 84*, 289-297.

Iyer, H. K., Berry, K. J., & Mielke, P. W. (1983). Computation of finite population parameters and approximate probability values for multi-response randomized block permutation procedures (MRBP). *Communications in Statistics – Simulation and Computation, 12*, 479-499.

Janson, H., & Olsson, U. (2001). A measure of agreement for interval or nominal multivariate observations. *Educational and Psychological Measurement, 61*, 277-289.

Krippendorff, K. (1970). Bivariate agreement coefficients for reliability of data. *Sociological Methodology, 2*, 139-150.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Newbury Park, CA: Sage.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research, 28*, 587-604.

Mielke, P. W. (1979). On asymptotic non-normality of null distributions of MRPP statistics. *Communications in Statistics – Theory and Methods, A8*, 1541-1550.

Mielke, P. W. (1984). Meteorological applications of permutation techniques based on distance functions. In P. R. Krishnaiah and P. K. Sen (Eds.), *Handbook of statistics, volume 4* (pp. 813-830). Amsterdam: North-Holland.

Mielke, P. W., & Berry, K. J. (1982). An extended class of permutation techniques for matched pairs. *Communications in Statistics – Theory and Methods, 11*, 1197-1207.

Mielke, P. W., & Berry, K. J. (1994). Permutation tests for common locations among samples with unequal variances. *Journal of Educational and Behavioral Statistics, 19*, 217-236.

Mielke, P. W., & Berry, K. J. (1999). Multivariate tests for correlated data in completely randomized designs. *Journal of Educational and Behavioral Statistics, 24*, 109-131.

Mielke, P. W., & Berry, K. J. (2001). *Permutation methods: A distance function approach*. New York, NY: Springer-Verlag.

Mielke, P. W., & Iyer, H. K. (1982). Permutation techniques for analyzing multi-response data from randomized block experiments. *Communications in Statistics – Theory and Methods, 11*, 1427-1437.

Mielke, P. W., & Sen, P. K. (1981). On asymptotic non-normal null distributions for locally most powerful rank test statistics. *Communications in Statistics – Theory and Methods, A10*, 1079-1094.

Mielke, P. W., & Yao, Y. C. (1990). On g-sample empirical coverage tests: Exact and simulated null distributions of test statistics with small and moderate sample sizes. *Journal of Statistical Computation and Simulation, 35*, 31-39.

Mielke, P. W., Berry, K. J., & Johnson, E. S. (1976). Multi-response permutation procedures for *a priori* classifications. *Communications in Statistics – Theory and Methods, A5*, 1409-1424.

Mielke, P. W., Berry, K. J., Brockwell, P. J., & Williams, J. S. (1981). A class of nonparametric tests based on multiresponse permutation procedures. *Biometrika, 68*, 720-724.

Moffett, E. A., & Dominick, J. R. (1987). Statistical analysis in the JOB 1970-85: An update. *Feedback, 28*, 13-16.

O'Reilly, F. J., & Mielke, P. W. (1980). Asymptotic normality of MRPP statistics from invariance principles of *U*-statistics. *Communications in Statistics – Theory and Methods, A9*, 629-637.

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research, 27*, 258-284.

Riffe, D., & Freitag, A. A. (1998). A content analysis of content analysis: Twenty-five years of Journalism Quarterly. *Journalism and Mass Communication Quarterly, 74*, 873-882.

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly, 19*, 321-325.

Stewart, T. D. (2002). *Principles of research in communication*. Boston: Allyn & Bacon.

Tucker, D. F., Mielke, P. W., & Reiter, E. R. (1989). The verification of numerical methods with multivariate randomized block permutation procedures. *Meteorology and Atmospheric Physics, 40*, 181-188.

Wilhoit, F. G. (1984). Student research productivity: Analysis of Journalism Abstracts. *Journalism Quarterly, 61*, 655-661.

Wimmer, R. D., & Dominick, J. R. (1994). *Mass media research: An introduction* (4th ed.). Belmont, CA: Wadsworth.

Zimmerman, G. M., Goetz, H., & Mielke, P. W. (1985). Use of an improved statistical method for group comparisons to study effects of prairie fire. *Ecology, 66*, 606-611.

Appendix I

Permutation tests represent the "ideal" situations where one can derive the exact probabilities rather than approximate values obtained from common probability distributions, such as the t, F and χ^2 (Mielke & Berry, 2001, p. 1). Carrying out randomization or permutation of the collected data rather than relying on the often times unreasonable assumption of random sampling or normality not only make a test more data-dependent, but also enhances the practicability of a test, as the practitioners have full control over the stochastic component of the statistical model.

Subsequent development of this procedure includes: (1) the asymptotic behavior of the null distributions of the MRPP statistics (Mielke, 1979; O'Reilly & Mielke, 1980; Mielke & Sen, 1981), (2) specifications of the MRPP model when analyzing univariate rank data (Mielke et al., 1981), (3) appropriate techniques for analyzing Multi-Response data from Randomized Blocks layout (MRBP) (Mielke & Iyer, 1982), (4) a class of techniques for matched-pairs t test (Mielke & Berry, 1982), (5) computational procedures for finite population parameters of MRPP and MRBP (Berry & Mielke, 1983a; Iyer, Berry, & Mielke, 1983), (6) moment approximations as an alternative to the F test for detection of location and scale shifts under variance heterogeneity (Berry & Mielke, 1983b; Mielke & Berry, 1994), (7) generalization of Cohen's κ to multiple coders and higher levels of measurement (Berry & Mielke, 1988), (8) goodness-of-fit empirical coverage tests (Mielke & Yao, 1990), (9) multivariate measures of association for nominal and higher levels of variables (Berry & Mielke, 1992), (10) intercoder agreement measure comparisons between two independent sets of coders (Berry & Mielke, 1997a), (11) measuring the agreement between coders and a standard (Berry & Mielke, 1997b), (12) multivariate tests for correlated dependent variables in randomized experiments (Mielke & Berry, 1999). This method has been successfully applied in the fields of meteorology and ecology (Mielke, 1984;

Tucker, Mielke, & Reiter, 1989; Zimmerman, Goetz, & Mielke, 1985). In addition, least sum of absolute deviations regression, coupled with MRPP yields a versatile and robust linear model that is capable of handling data from complex factorial experimental designs with covariates (Mielke & Berry, 2001).

Assuming an *r*-dimensional *k* group design with the combined sample size equaling *N*, and group sizes equaling n_i where $i = 1, \ldots, k$, and $\sum_{i=1}^{k} n_i = N$, let $(x_{1I}, \ldots, x_{rI})$ denote the r-dimensional responses where $I = 1, \ldots, N$, and let S_i, where $i = 1, \ldots, k$ denote the *k* groups of responses, or using the terms of Mielke and Berry (2001), the "exhaustive partitioning" of the *N* responses into *k* disjoint sets (p. 12). The basic formulation of the MRPP family of statistics involves the definition of a symmetric distance function of the form

$$\Delta_{I,J} = \left[\sum_{c=1}^{r}(x_{cI} - x_{cJ})^2\right]^{v/2},$$

as a measure of the multivariate distance between the two observations x_I and x_J. For notational simplicity both the "excess group" and the truncation of distance to a preset maximum value shall not be discussed in the present paper (for details see Mielke & Berry, 2001). The choice of v is arbitrary, but the two choices $v = 1$ and $v = 2$ seems most reasonable. When $v = 1$, the distance is metric Euclidean distance and this distance function has nice theoretical properties of being robust and much less influenced by outliers (Mielke & Berry, 2001). When $v = 2$, the distance is defined in a non-metric squared Euclidean space because the triangle inequality fails in this analysis space, and it is known through both theoretical and simulative studies that this choice leads to a less robust test (Mielke & Berry, 1994). However, the choice of $v = 2$ yields an easier explanation of the test results, because many popular tests essentially involve the use of squared distance.

The MRPP statistic can be thought of as a weighted average of within-group distances. Intuitively, a smaller value of the MRPP statistic would mean higher concentration within each *a priori* classified group (Mielke, 1984, p. 815). Such an interpretation is also in line with the geometric interpretation of the conventional multivariate analysis of variance (see Edgington, 1987, pp. 190). Therefore, in terms of detecting between group differences, a smaller value of the MRPP statistic is necessarily "better."

The MRPP statistic is given by

$$\delta_{obs} = \sum_{i=1}^{k} C_i \xi_i ,$$

where C_i is the group weight for $i = 1, \ldots, k$, and $\sum_{i=1}^{k} C_i = 1$, and

$$\xi_i = \binom{n_i}{2}^{-1} \sum \Delta_{I,J}\, \psi(x_I)\, \psi(x_J) ,$$

is the average within-group distance for all distinct pairs of responses in the *i*th group. $\psi(\cdot)$ is an indicator function given by

$$\psi(x_I) = \begin{cases} 1 & \text{if } x_I \in S_i , \\ 0 & \text{if } x_I \notin S_i . \end{cases}$$

The choice of group weights is extensively discussed in Mielke (1984), but $C_i = n_i / N$, and $C_i = (n_i - 1) / (N - k)$ are two sensible choices for $v = 1$ and $v = 2$, respectively.

The formal test of significance of δ_{obs} is carried out by assuming the null hypothesis of equal probabilities being placed upon each one of the

$$M = \frac{N!}{\prod_{i=1}^{k} n_i} ,$$

possible permutations of the N responses into the k groups, each permutation yielding a realized value of δ. The probability value associated with δ_{obs} is a ratio of the number of δ s being smaller than or equal to δ_{obs} and M, formally written as $P(\delta_{obs}) = \{\# \; \delta \leq \delta_{obs}\} / M$.

Because M is usually a very large number even for relatively small sample sizes, the exact reference distribution of the MRPP statistic is difficult to obtain, therefore, Mielke, Berry and Johnson (1976) have provided efficient computational methods for the first three cumulants of the MRPP null distribution, upon which a moment approximation using Pearson type III distribution may be utilized. Generally this approximation is excellent. For details please refer to Mielke & Berry (2001).

Appendix II

The following is a GAUSS procedure that implements the missing-cell situation of continuous reliability data. The input is a *b*-coder by *g*-unit matrix similar to Table 4. The output is a 3-by-1 vector, call it *resultv*, with the first element being the observed disagreement, the second element being the expected disagreement (un-averaged), and the third element being the denominator of expected disagreement. So agreement is simply 1 – *resultv*[1]/(*resultv*[2]/ *resultv*[3]).

```
proc discrete(x,v);
  local M, delta, distance, temp, x1, x2, n1, n2, distance1, factor1, factor2, resultm;
  M = prodc(sumc((x .gt 0)')!);
  delta = 0;
  distance = 0;
  for idxI (1, rows(x)-1, 1);
    for idxJ (idxI+1, rows(x), 1);
      x1 = x[idxI,.]';
      x2 = x[idxJ,.]';
      n1 = sumc(x1 .gt 0);
      n2 = sumc(x2 .gt 0);
      if (n1 lt n2);
        temp = n1;
        n1 = n2;
        n2 = temp;
        temp = x1;
        x1 = x2;
        x2 = temp;
      endif;
      x1 = x1[1:n1];
      x2 = x2[1:n1];
      delta = delta + sumc(abs(x1-x2)^v);
      distance1 = 0;
      for i (1, n1, 1);
        for j (1, n2, 1);
          distance1 = distance1+abs(x1[i] - x2[j])^v;
        endfor;
      endfor;
      factor1 = M*(1/n1);
      factor2 = M*(1-n2/n1);
      distance = distance + distance1*factor1+sumc(x1^v)*factor2;
    endfor;
  endfor;
  resultm = zeros(3,1);
  resultm[1] = delta/(rows(x)!/(rows(x)-2)!/2);
  resultm[2] = distance/(rows(x)!/(rows(x)-2)!/2);
  resultm[3] = M;
  retp(resultm);
endp;
```

Appendix III

The following is a GAUSS procedure that implements the method based on cumulative lengths. The input is a *b*-coder by *g*-unit matrix similar to Table 5. The output is a 3-by-1 vector, call it *resultv*, with the first element being the observed disagreement, the second element being the expected disagreement (un-averaged), and the third element being the denominator of expected disagreement. So agreement is simply 1 – *resultv*[1]/(*resultv*[2]/ *resultv*[3]).

```
proc continuous(x);
  local delta, distance, M, temp, x1, x2, n1, n2, distm, tempdistm, part2,
        indexm, minor, minorv, factor, resultm;
  delta = 0; distance = 0;
  M = prodc(sumc((x .gt 0)')!);
  for idxI (1, rows(x)-1, 1);
    for idxJ (idxI+1, rows(x), 1);
      x1 = x[idxI,.]';
      x2 = x[idxJ,.]';
      n1 = sumc(x1 .gt 0);
      n2 = sumc(x2 .gt 0);
      if (n1 lt n2);
        temp = n1;
        n1 = n2;
        n2 = temp;
        temp = x1;
        x1 = x2;
        x2 = temp;
      endif;
      x1 = x1[1:n1];
      x2 = x2[1:n1];
      for i (1, n1, 1);
        delta = delta + (sumc(x1[1:i])-sumc(x2[1:i]))^2;
      endfor;
      distm = zeros(n1, n1);
      for i (1, n1, 1);
        distm[.,i] = x1[.,1] - x2[i,1];
      endfor;
      for j (1, n2, 1);
        distance = distance+sumc(vec(distm[.,1:n2]^2))*(1/n1/n2)*M*(n1+1-j);
      endfor;
      if (n2 ge 2);
        for i (1, n2-1, 1);
          for j (i+1, n2, 1);
            tempdistm = distm[.,1:n2];
            for col (1, n2-1, 1);
              for row (1, n1, 1);
                indexm = zeros(n1,n2);
                indexm[row,.] = ones(1,n2);
                minor = delif(tempdistm,indexm);
                minorv = tempdistm[row,col].*vec(minor[.,col+1:n2]);
                factor = 4*(1/n1)*(1/n2)*(((n2-1)*(n1-1))^(-1))*M*(n1+1-j);
                distance = distance + factor*sumc(minorv);
              endfor;
            endfor;
          endfor;
        endfor;
      endif;
      if ((n1 - n2) ge 1);
        for j (n2+1, n1, 1);
          distance = distance + sumc(distm[.,j]^2)*(1/n1)*M*(n1+1-j);
        endfor;
        for i (1, n2, 1);
          for j (n2+1, n1, 1);
            tempdistm = distm[.,1:n2];
            part2 = distm[.,j];
            for col (1, n2, 1);
              for row (1, n1, 1);
                indexm = zeros(n1,1);
                indexm[row] = 1;
                minor = delif(part2,indexm);
                minorv = tempdistm[row, col].*minor;
                factor = 2*(1/n1)*(1/n2)*((n1-1)^(-1))*M*(n1+1-j);
                distance = distance + factor*sumc(minorv);
              endfor;
            endfor;
          endfor;
        endfor;
        for i (n2+1, n1-1, 1);
          for j (i+1, n1, 1);
            tempdistm = distm[.,i];
            part2 = distm[.,j];
            for row (1, n1, 1);
              indexm = zeros(n1,1);
              indexm[row] = 1;
              minor = delif(part2,indexm);
              minorv = tempdistm[row].*minor;
              factor = 2*(1/n1)*((n1-1)^(-1))*M*(n1+1-j);
              distance = distance + factor*sumc(minorv);
            endfor;
          endfor;
        endfor;
      endif;
    endfor;
  endfor;
  resultm = zeros(3,1);
  resultm[1] = delta/(rows(x)!/(rows(x)-2)!/2);
  resultm[2] = distance/(rows(x)!/(rows(x)-2)!/2);
  resultm[3] = M;
  retp(resultm);
endp;
```

CT&M

How General Principles of Organization Theory Explain Gatekeeping Decisions About News: A Revised View of the Field

by Hugh J. Martin
Assistant Professor
The University of Georgia
Grady College of Journalism and Mass Communication
Athens, Ga. 30602-3018
Phone: (706) 542-5033
Fax: (706) 542-2183
email: hjmartin@arches.uga.edu

A paper presented to the Communication Theory and Methodology Division of the Association for Education in Journalism and Mass Communication, Annual Convention, Kansas City, MO, July 30-August 2, 2003.

Abstract

Mass communication hypotheses about gatekeeping do not provide a coherent explanation of organizational influences. General theories define organizations as collections of individuals working toward a common goal. This view focuses on the processes that organizations develop, and the internal and external variables that influence the outcome of those processes. These principles are used to develop alternative models of news selection. New propositions about news selection are derived from these models.

Introduction

Mass communication theorists state the conceptualization of gatekeeping needs to be improved (Hirsch, 1977; Shoemaker, 1991). These theorists state analysis of gatekeeping behavior from a high level of theoretical abstraction will better explain how news is selected.

This study responds to those suggestions by drawing from general theories that describe organizations and their behavior (Adams, 1980; Barnard, 1968; Donaldson, 1996: Galbraith, 1973; March & Simon, 1958; Pfeffer & Salancik, 1978; Thompson, 1967). Gatekeeping is reconceptualized by identifying important characteristics of organizations that may influence decisions about the publication of news. Alternative models of gatekeeping are developed. The study describes how these theoretical models can be used to predict dynamic interactions that shape the production of news. The study also makes suggestions for additional research to test these models.

Justification for the Study

Gatekeeping studies belong to a four-decade line of research that examines how information is channeled to news organizations, selected for use, and then reported as news (Shoemaker, 1991). Gatekeeping originated with Kurt Lewin's model identifying the "gates" that food passed through on its way from store or garden to the dinner table (Shoemaker, 1991, p. 5-9). Lewin argued that gatekeeping applied generally to social processes that use channels to move objects past a series of decision points (p.9).

Mass communication researchers adopted gatekeeping to examine "the process by which the billions of messages that are available in the world get cut down and transformed into the hundreds of messages that reach a given person on a given day" (Shoemaker, 1991, p.1). Gatekeeping studies focus on the rules and forces that determine whether information gathered by news organizations is allowed to pass critical decision points. By comparing information available to news organizations with the information that is published, gatekeeping provides "a conceptual structure for comparing media content with some other measure of 'reality'" (Shoemaker, 1991, p 3.)

However, some gatekeeping studies cannot explain why news decisions were made. The first gatekeeping study (White, 1950) attempted to describe how a newspaper wire editor's "highly subjective ... value judgments" (p. 386) guided news selection. As Hirsch (1977, p. 22-23) subsequently pointed out, the wire editor's selections almost mirrored the distribution of topics received from wire services.

Whitney and Becker (1982), responding to similar results in several studies, conducted an experiment manipulating the distribution of topics offered to newspaper and television wire editors. Results supported a hypothesis that the categories of stories "transmitted by wire services served to cue editors as to proportions [of topics] which should be selected by those editors" (p. 65).

Wilke and Rosenberger (1994) found that Associated Press editors in Germany, who translated news originating in the United States, also made selections matching the distribution of topics and news content in the original universe of stories. "The similarity of input and output is contrary to our expectation that news agency editors would follow certain patterns when determining which information is important and interesting to the German media," (p. 423) the researchers wrote.

Hirsch (1977) suggested such results stem from not separating the unique social functions of news from the organizational processes that are used to produce news. He argued that the behavior of news organizations, and individuals within them, can best be explained with general principles that apply to the study of organizations.

Shoemaker (1991, p. 76-77) identified three problems with gatekeeping studies. First, results are confusing when variables from different levels of analysis are used in the same study. Second, there is an inadequate understanding of links between different levels of analysis, such as individual and organizational influences. Third, there is a failure to distinguish the relative importance of various influences on gatekeeping decisions. Shoemaker stated the theoretical basis of gatekeeping studies should be extended to "higher levels of analysis" (p. 76) to resolve these difficulties. She cited an article (Adams, 1980) from organization theory as an example.

This study focuses on organizational influences in response to these suggestions. The study develops general descriptions of the characteristics that define news organizations, of how news organizations process information, and of how news organizations interact with their environments. These descriptions are then used to develop empirically testable propositions describing how organizational processes influence the selection of news.

Defining News Organizations

Gatekeeping researchers have long been interested in organizational variables (see, e.g., Abbott & Brassfield, 1989; Bantz, 1990; Bass, 1969; Beam, 1996; Demers, 1996; Gieber & Johnson, 1961; Tuchman, 1972; Turow, 1990; Whitney, 1981). However, gatekeeping studies often fail to distinguish organizational influences from other variables (Abbott & Brassfield, 1989, p. 855; Harless, 1974, p. 88; Snider, 1967, p. 420).

Organizational influences can be identified by first defining the general characteristics of organizations. These characteristics can then be applied to the description of news organizations.

Barnard (1968) defined an organization as individuals working together to reach a goal. The coordination of individual efforts overcomes the physical and psychological limitations faced by individuals working alone.

The most basic goal of any organization is survival. However, as Barnard (1968) pointed out, individual members are not always committed to organizational goals. March and Simon (1958) described how organizations try to gain cooperation from individuals by offering incentives to participate that are greater than the cost of participation. Even with incentives, there is potential for internal disagreements about organizational goals. Organizations manage such conflict by developing programs, or routines, that define and coordinate desirable responses to a given stimulus (March & Simon,1958). Organizational routines compensate for the cognitive limits that individuals have when trying to find appropriate responses to a given stimulus. Routines also provide a stable framework that constrains and guides decision-making (March & Simon, 1958).

The assertion by March and Simon (1958) that routines are a fundamental characteristic of organizations is important because Shoemaker's (1991) analysis of gatekeeping separates routines from organizations. Shoemaker (1991, chap. 3) identified five levels of analysis in gatekeeping research as (1) individual decision making, (2) communication routines, (3) organizational influences, (4) institutional forces, and (5) social forces. Shoemaker identified "communication routines" (p. 48) as patterned, repeated practices used to gather, process and transmit the news and to communicate about these tasks. Shoemaker defined routines to "include communication practices that are common across many communication organizations, where we reserve the organizational level for those factors on which communication practices may vary" (p. 53).

This separation misstates the nature of organizations. First, routines are one of the defining characteristics of organizations (March & Simon, 1958). Second, organizational theorists (Barnard, 1968; Galbraith, 1973; March & Simon, 1958; Pfeffer & Salancik, 1978; Thompson, 1967) specify that organizational behavior can be systematically explained by examining the interactions between key variables both within an organization and in the organization's environment. These interactions are defined by organizational routines. Therefore, organizational characteristics cannot be separated from organizational routines.

To explain routines as an extra-organizational phenomenon, Shoemaker (1991) cited a classic case study of routines that may exist in many news organizations. The study (Tuchman, 1972) of a single newspaper concluded that the procedures used to create news "*are actually strategies through which newsmen protect themselves from* [internal] *critics and lay professional claim to objectivity* [italics in original]..."p. 676.[1]

However, organization theory (March & Simon, 1958) suggests that news organizations which want to be considered objective will create routines to reach that goal. Therefore, all news organizations that have a goal of claiming objectivity are likely to create similar routines.

Structuring News Organizations

Given the nature of organizations -- collections of individuals who are persuaded to engage in coordinated routines designed to reach organizational goals -- how is their structure determined? Organization theorists have identified three major influences on structure.

---

[1] Procedures were (a) presenting conflicting claims about the truth, (b) using supplemental evidence to support facts, (c) using quote marks to show a reporter did not make a statement, (d) presenting the most relevant facts first, and (e) labeling stories containing "opinion" as news analysis (Tuchman, 1972, p. 676).

The effects of size

Size is the first influence. Organizations require resources to reach their goals (Pfeffer & Salancik, 1978; Thompson, 1967). Donaldson (1996, p. 169) argued that organizations make incremental adaptations to contingencies in their environments so they can acquire more resources and increase their effectiveness. As organizations grow, they become more differentiated -- the number of hierarchical levels increases along with the number of departments managing the division of labor at each level (p. 166). Specialization and bureaucratization, or the use of standardized policies and administrative procedures, also increase with size. However, decision making becomes more decentralized (Donaldson, 1996).

There is some evidence that gatekeeping processes become more complex, involving more decision makers, as news organizations increase in size (see, e.g. Abbott & Brassfield, 1989, p. 855; Berkowitz, 1990, p. 62). However, studies did not control for size differences. Some research argued that increases in the complexity of gatekeeping decisions are associated with the type of news organization, such as television stations (Berkowitz, 1990). More research is needed in this area.

This discussion suggests the first set of propositions:

1a. The complexity of news routines is positively related to organizational size. Larger news organizations involve more people in gatekeeping decisions.

1b. The use of policies, written and unwritten, to control gatekeeping decisions is positively related to organizational size.

1c. The influence of policies will be moderated by the level where the policy exists because gatekeeping decisions will be more decentralized as news organizations increase in size.

1d. The previous proposition suggests there will be a positive relationship between organizational size and the influence of individuals on gatekeeping decisions that are subject to policy directives from managers at the top of the hierarchy.

Mechanical vs. organic organizations

Durkheim's (see Morrison, 1995, p. 129-130) concept of mechanical and organic societies is the second major influence on organizational structure. Mechanical societies have a collective identity that supersedes individual autonomy. Mechanical societies use common beliefs and practices to mobilize collective action. These societies tend to be smaller and more homogenous, lacking specialized divisions of labor. Organic societies, however, divide labor. In organic societies individuals rely on others to perform functions that they cannot carry out. Individuals in organic social structures are more closely linked to each other than to the collective society.

Organization theorists adopted this idea.[2] March and Simon (1958, p.147) stated there is little individual latitude in organizational routines that specify the performance of tasks contributing to an outcome. This description is clearly mechanical. Organic routines give individuals more latitude because they specify a desired outcome, but do not specify how to achieve the outcome. Donaldson (1996, p. 167) stated tasks which require innovation or involve a high degree of uncertainty require a more organic structure. Better educated, more professional workers are assigned to such tasks.

Shoemaker (1991, p. 49) argued that as routines become more important, the selection of news will become more uniform. However, organizational theory suggests that individual influence on the performance of organizational tasks can only be identified if researchers first identify a relevant routine as mechanical or organic. For example, editors who prepare for publication stories that have been through several drafts are performing a mechanical task by contributing to an organizational outcome. This may explain the lack of individual influence that some studies found (White, 1950;Whitney & Becker, 1982; Wilke and Rosenberger, 1994). A reporter's job, however, is organic because reporting routines focus on outcomes -- completing news stories. Reporters may exert more individual influence on how news is gathered and shaped. Therefore, gatekeeping studies must account for the nature of the task.

---

[2] Of course, societies are organizations too.

This suggests a second set of propositions:

2a. Individual influence on gatekeeping decisions is negatively related to the degree of control imposed by news processing routines. However, this relationship is moderated by two other relationships.

2b. There is a positive relationship between organic, or outcome-based, tasks and individual influence on gatekeeping decisions.

2c. There is a negative relationship between mechanical, or news processing, tasks and individual influence on gatekeeping decisions.

Processing information

Information processing requirements are the third major influence on organizational structure. Galbraith (1973) described how organizational structure is affected by uncertainty about how to complete tasks. Uncertainty (p. 5) is the difference between the amount of information required to perform a task and the amount of information the organization has. Information requirements are determined by (a) the diversity of the organization's output, (b) the number of different inputs the organization uses, and (c) how difficult it is to perform a task.

Tasks within organizations vary, so organizations face different degrees of uncertainty (Galbraith, 1973, p. 3). For example, if routines are a series of steps performed by individuals or groups, the uncertainty associated with each step, and its structure, may vary. This is another reason – in addition to organizational size and organic or mechanical routines – that different tasks must be structured differently. Different tasks must be structured differently so they can be integrated to achieve desired outcomes.

Managers, however, find it harder to integrate activities in larger organizations because they cannot speak to everyone face-to-face. Galbraith (p. 10) suggested organizational routines reduce the need for direct communication between managers and workers. However, some situations that are not accounted for by existing routines must still be referred to a manager. The number of exceptions to routines increases as tasks become more uncertain, and eventually managers cannot cope with all of the exceptions.

Managers can avoid this kind of overload either by reducing the need to process information, or by increasing the organization's capacity for processing information (Galbraith, 1973, chap. 3-5). The need to process information can be reduced by (a) reducing the level of performance to create slack resources, or (b) creating self-contained divisions within organizations to reduce the diversity of information that managers face. If organizations instead wish to increase the capacity for processing information they can (a) invest in systems for distributing information up the hierarchy, or (b) decentralize decision making so managers can resolve problems at the level where they occur.

Galbraith's (1973) argument suggests organizational structure will influence the diversity and complexity of news that organizations produce. For example, news organizations could reduce information processing requirements, creating slack resources, if they publish fewer stories that require complex sourcing or editing. News organizations could also reduce the amount of information processed by creating self-contained divisions to handle various routines.

News organizations that want to increase the diversity of news might instead increase information processing capacity using technology that increases internal distribution of information. Alternatively, such organizations could give lower ranking editors or reporters more authority over news decisions.

Two studies are consistent with Galbraith's (1973) theory, although neither cited his ideas. The first (Whitney, 1981) compared news processing at a wire service bureau and a radio station. The bureau (p. 72-74) pressed an editor and a salesman into service as reporters when breaking news overwhelmed available resources. However, the bureau did not spend more time searching for news on slower days. This may indicate the bureau did not have slack resources, but managers did have authority to redirect resources when news processing was overloaded.

Whitney (1981, p. 75-76) also found the radio station devoted less time to processing news on the slowest and on the busiest days. Story preparation time was longest on days with a moderate workload. This may indicate the station had slack resources which could be pressed into service, but once the slack was used up the station compensated by reducing the level of performance.

The second study (Demers, 1996) found a positive relationship between the complexity[3] of a newspaper's internal structure and the diversity of content published on its editorial pages. Recall that authority becomes decentralized in larger organizations (Donaldson, 1996). Therefore, Demers (1996) finding that larger organizations produce more diverse editorial pages may result from the increase in information processing capacity that Galbraith (1973) said results from decentralizing authority.

---

[3] Structural complexity was measured by the number of levels in the newspaper's hierarchy, the number of separate beats employing reporters, and the number of employees (Demers, 1996, p. 866).

Galbraith's (1973) arguments complement and extend the earlier discussion of organizational definitions and structure. For instance, Tuchman (1972) found news processing routines that were designed to avoid libel suits. Such routines would help meet an organizational goal of avoiding lawsuits, while increasing the news organization's information processing capacity by eliminating the need to refer potential legal problems to a manager.

This suggests a third set of propositions:

3a. The diversity of news covered and stories told is a function of organizational structure. Diversity is positively associated with the availability of slack, or unused, resources.

3b. The diversity of news covered and stories told is positively associated with the increased separation of news operations into different divisions.

3c. The association between diversity of news reported and an organization's investment in systems for processing information is indeterminate. Such systems give senior managers more control over the production of news, so their views are likely to have more influence.

3d. The diversity of news gathered and stories told is positively associated with the decentralization of authority to deal with exceptions to routines within news organizations.

Environmental Influences on News Organization

The discussion so far has focused on the internal characteristics of news organizations. However, the production of news requires information from sources outside the organization, and the resulting stories are distributed to an extra-organizational audience.

Mass communication scholars have long recognized that sources of information can influence news decisions. For example, Gandy (1982) argued that sources influence news decisions by making information available to journalists. This, in effect, reduces the cost of information by reducing the work needed to produce a story. Sources also can withhold information, increasing its cost and decreasing the likelihood that it will be published. Powers and Fico (1994, p. 88-89) discuss research examining the influence of sources on news organizations. Turow (1990) used an organizational model to explore how the medical profession influenced the content of television shows featuring doctors.

This suggests the importance of a theory explaining how internal processes are affected by an organization's environment. Thompson (1967) argued that organizations attempt to gain control of their environment through growth. Growth allows organizations to absorb critical resources. However, there are limits to an organization's ability to absorb its environment -- some critical resources will always be controlled by other organizations. Pfeffer and Salancik (1978) described how dependence on external resources can result in the external control of an organization.

The resource-dependence model provides a direct link between the organization's environment and its internal routines. Pfeffer and Salancik (1978, p. 32) defined an organization by its behavior, so individual membership is defined by the extent of participation in organizational behavior. The key is the degree of control that the organization has over an activity. The organization's boundary is found where its ability to control activities becomes less than the ability of another organization or individual to control that same activity. For example, a reporter gathering information from sources is a member of a news organization, but that same reporter eating dinner after work is not.

Conditions moderating control of the news

Once an organizational boundary is crossed, resources are controlled by other organizations or individuals. These other organizations or individuals *might* determine whether resources are provided to the organization, and under what conditions. However, the influence of these external actors is moderated by five key conditions (Pfeffer & Salancik, 1978, p. 44). First, the external actor must control access to a resource that is critical to an organization's survival -- there cannot be alternative resources available to the organization. Second, the organization cannot also control access to a resource that is critical to the external actor. Third, the organization's outputs must be visible to the external actor so that it can judge compliance with its demands. Fourth, the organization cannot control the determination, formulation or expression of the external actor's demands. Fifth, the organization must be capable of developing outcomes that satisfy the external actor's demands.

The extent to which each of these conditions holds for news organizations will determine the influence of external actors – such as sources of information – on the production of news.

The first condition would require that sources control information that is critical to a news organization's survival. Pfeffer and Salancik (1978) argued dependence on a resource is determined both by the importance of the resource and by "the extent to which it is controlled by a relatively few" (p. 51) external organizations.

This suggests that as the number of alternative sources of information increases, the influence of any single source can exercise will decrease. For example, news organizations might be expected to report uncritically on leaks of secret investigative information when an investigator is the only source of that information. However, news organizations will have more control over coverage of controversial government proposals once those proposals are made public because an array of sources can then provide information.

The second condition requires that news organizations cannot control access to a resource that sources consider important. Just the opposite may be the case in markets with few news organizations. News organizations control access to their audiences by determining what is published. Sources cannot access the organization's audience unless the news organization chooses to publish the sources' information.

The third condition requires that sources can observe stories which are published to determine whether information they provide is included. Many sources probably can easily obtain, view, or listen to stories news organizations produce, and therefore observe whether organizations comply with their demands.[4]

The fourth condition would require that sources control access to information that news organizations need, and can therefore set the conditions of access to that information. However, controlling access and setting conditions requires possession of a resource (Pfeffer & Salancik, 1978, p. 48). In the United States, much information about the government is considered public under the First Amendment to the Constitution. Laws often specify which government meetings and records should be open to public scrutiny, and which should not. Therefore, the law limits the government's ability to control access to information. News organizations have a long history of litigation to preserve such laws, and this can be viewed as an effort to maintain control over access to an important resource.

However, such laws often apply only indirectly to private organizations when those organizations are required to provide information to the government. Therefore, private organizations may have much greater control over the distribution of information in their possession. For example, sources on both sides of a local controversy told researchers that they had limited access by reporters to private meetings, concealing important information about the issue (Fico & Soffin, 1994, p.71).

---

[4] This may not be true if the source lives outside the circulation area or range of a broadcast signal.

The fifth condition would require that news organizations shape their reports to include or exclude information in response to suggestions or demands made by sources. News organizations could easily do this if they wished.

The resource-dependence model suggests, then, that gatekeeping studies must account for factors influencing the control of information when examining external influences on how news is gathered and reported. The model also suggests that organizational preferences for the type and volume of news coverage will be influenced by the degree of control over information in the hands of news sources, and also by the desire of those sources to provide information to the news organization's audience. This leads to the fourth set of propositions:

4a. There will be a negative relationship between the number of alternative sources of information and source control over the resulting stories.

4b. Source control over stories will be moderated by the assignment of rights to information. As laws require the disclosure of more information, source control will diminish.

Extending the model

Adams (1980) extended the resource-dependence model (Pfeffer & Salancik, 1978) to examine how so-called boundary role persons behave. Adams (1980, p. 328) stated boundary-role persons acquire and filter organizational inputs and outputs, collect information, represent the organization to its environment, and buffer the organization from external threats and pressures. This description fits well with the activities of reporters and others at news organizations who are in regular contact with other organizations.

Adams (1980, p. 333-335) stated boundary role persons who negotiate the acquisition of resources and disposal of outputs face an inherent conflict -- they must follow organizational directives while accurately and influentially representing the preferences of the resource supplier to the organization. This, however, can raise suspicions about the boundary role person's loyalty. This distrust can force the boundary role person to bargain more competitively with the resource supplier, risking an unsuccessful negotiation. This, in turn, would breed more distrust of the boundary-role person within the organization.

Adams (1980, p. 337-338) also suggested that organizations can most easily detect errors when boundary role persons either admit the wrong input or produce the wrong output. If a boundary role person fails to admit a correct input, that error is least likely to be detected by the organization. Detection of failure to produce a correct output will depend on feedback that the organization receives.

This suggests reporters may reduce the potential for conflict with their organizations by withholding information that conflicts with organizational directives. News organizations may unintentionally reinforce this behavior. This is because if reporters provide information that is not considered acceptable, it will be rejected. However, reporters may extrapolate that similar information is likely to be rejected, and withhold that too. This may explain why one study (Gieber, 1960, p. 203) found that reporters were making news decisions based on policies which their news organizations did not actually have.

However, organizations may discover reporters are withholding information if there is competition from another news organization covering the same story. If a competitor publishes information that an organization did not publish, the organization is alerted to the fact that information was withheld. For example, Beam's (1996) study of editor perceptions found the presence of a competing newspaper is the best predictor of uncertainty among editors about what information their readers demand.

This suggests a fifth set of propositions:

5a. Source influence will be positively related to requests to withhold information.

5b. Source influence will be negatively related to requests to publish information.

5c. Source influence will be moderated by competition. As competition increases, requests to withhold information become less influential because reporter errors become increasingly visible.

5d. As competition increases, source requests to publish information become more influential because news organizations are increasingly uncertain about what audiences want.

Discussion

The organizational models developed here are summarized in Table 1. These models allow researchers to better identify how organizational characteristics influence news processing, thereby separating those influences from what Hirsch (1977) called the social functions of news.

Table 1 Goes About Here

The models also respond to Shoemakers' (1991) three concerns about inadequate conceptualization of gatekeeping. First, they specify variables that belong at the organizational level of analysis. Table 1 shows these include (a) organizational size, (b) the nature of organizational routines, (c) how organizations process information, (d) characteristics of an organization's environment, and (e) interactions between organizational imperatives and demands from news sources.

Second, the models clarify links between individual and organizational influences. For example, Table 1 shows organizations decentralize as their size increases, which allows individuals more influence over decisions. Table 1 also shows the nature of organizational routines helps determine the influence of individuals.

Third, the models help distinguish the importance of different influences on gatekeeping decisions. For example, Table 1 shows organizations must have resources to process additional information if they want to increase the diversity of their news report. Table 1 also shows legal requirements are an important variable that determines what information is available to news organizations. Both examples suggest variables – slack resources, existing laws – that when present can be expected to account for major portions of the variance in gatekeeping decisions.

Consider two specific examples. First, there is evidence wire editors select stories that mirroring the distribution of topics they receive (White, 1950;Whitney & Becker, 1982; Wilke and Rosenberger, 1994). This is consistent with the proposition in Table 1 predicting a negative relationship between news-processing tasks and individual influence. This suggests a hypothesis that wire editors are primarily engaged in routine news processing, and therefore make few decisions that change the distribution of news they receive. This hypothesis could be tested further by studies examining whether editors at higher levels whose jobs focus on outcomes rather than processing actually determine the distribution of stories that are published.

A second example is Whitney's (1981) study reporting news organizations altered their capacity to produce news as the number of stories they processed varied. This is consistent with the proposition in Table 1 predicting story diversity is positively associated with slack resources. This suggests a hypothesis that the number of newsroom employees available to cover and process news is positively associated with the range of topics covered by an organization.

More generally, Table 1 shows how hypotheses can be developed for each of the five major organizational variables in the first column. One hypotheses is that organizational size is positively related to the number of people involved in a gatekeeping decision. A second hypotheses is that size is positively related to the number of divisions where individual gatekeeping decisions are made.

Routines are the second major variable. Table 1 suggests a hypothesis that individual influence on gatekeeping decisions is positively associated with jobs that involve either gathering news, or directing newsgathering.

The third major variable is capacity to process information. Table 1 suggests a hypothesis that the number of news departments within an organization is positively related to the range of topics covered by the organization.

The fourth variable is environmental conditions. Table 1 suggests a hypothesis that there is a positive association between laws requiring disclosure of information and publication of that information in news stories.

The fifth variable is the relationship between sources and reporters. Table 1 suggests a hypothesis that competition among news organizations is positively related to the number of stories that are critical of sources who often provide reporters with information about the news.

Conclusion

Gatekeeping takes place within news organizations, so researchers must account for organizational variables to understand gatekeeping behavior. This study develops models of organizational influences on gatekeeping. The models define organizations as groups of people who coordinate their activities in pursuit of some goal -- the most basic organizational goal being survival. The interdependence that characterizes organizations suggests that in addition to outcomes, research should focus on the behavior of organizations as manifested in joint actions, or processes (Pfeffer & Salancik, 1978, p. 40).

Studies that do not fully explicate organizational variables are unlikely to produce coherent explanations of gatekeeping behavior. Even if organizational variables are included, their theoretical implications may be misunderstood if they are not derived from a more general theory of organizations. Reynolds (1971, p.140-141) warned it is exceptionally difficult to develop theories about complex social phenomena by examining processes, then searching for patterns in the resulting data. However, some gatekeeping researchers seem to be using just this kind of bootstrap approach. Using a well developed general theory to develop empirically testable hypotheses is a more efficient approach.

Table 1: Summary of organizational influences on gatekeeping decisions

Proposition	1	2	3	4
Major variable				
Size	No. people involved	Policies controlling decisions	Decentralization of decisions [a]	Individual influence on high-level directives [a]
Relationship	+	+	+	+
Routines	Individual influence	Outcome-based task, individual influence	News processing task, individual influence	
Relationship	–	+	–	
Information Processing	Story diversity, slack resources	Story diversity, no. of news divisions	Story diversity, technology	Story diversity, decentralization
Relationship	+	+	no prediction	+
Environment's influence	No. alternative sources for stories	Legal requirements to disclose information		
Relationship	–	–		
Influence of sources	Requests to withhold information	Requests to publish information	Increased competition, requests to withhold [b]	Increased competition, requests to publish [b]
Relationship	+	–	–	+

[a] Moderates the influence of policy directives.

[b] Moderates the influence of source requests.

References

Abbott, E. A., & Brassfield, L. T. (1989). Comparing decisions on releases by TV and newspaper gatekeepers. *Journalism Quarterly, 66*(4), 853-856.

Adams, J. S. (1980). Interorganizational processes and organization boundary activities. In B. M. Staw & L. L. Cummings (Eds.), *Research in Organizational Behavior* (Vol. 2, pp. 321-355). Greenwich, CONN: Jai Press Inc.

Bantz, C. R. (1990). Organizational communication, media industries, and mass communication. In J. A. Anderson (Ed.), *Communication Yearbook* (Vol. 13, pp. 502-510). Newbury Park: Sage.

Barnard, C. I. (1968). *The Functions of the Executive*. (30th anniversary ed.). Cambridge, MASS: Harvard University Press.

Bass, A. Z. (1969). Refining the "gatekeeper" concept: A UN radio case study. *Journalism Quarterly, 46*(1), 69-72.

Beam, R. A. (1996). How perceived environmental uncertainty influences the marketing orientation of U.S. daily newspapers. *Journalism & Mass Communication Quarterly, 73*(2), 285-303.

Berkowitz, D. (1990). Refining the gatekeeping metaphor for local television news. *Journal of Broadcasting & Electronic Media, 34*(1), 55-68.

Demers, D. P. (1996). Corporate newspaper structure, editorial vigor, and social change. *Journalism & Mass Communication Quarterly, 73*(4), 857-877.

Donaldson, L. (1996). *For Positivist Organization Theory*. London: Sage.

Fico, F., & Soffin, S. (1994). Covering local conflict: Fairness in reporting a public policy issue. *Newspaper Research Journal, 15*(4), 64-76.

Galbraith, J. R. (1973). *Designing complex organizations*. Reading, MASS: Addison-Wesley Publishing Co.

Gandy, O. O. (1982). *Beyond Agenda Setting: Information Subsidies and Public Policy*. Norwood, NJ: Ablex Publishing.

Gieber, W. (1960). How the "gatekeepers" view local civil liberties news. *Journalism Quarterly, 37*(spring), 199-205.

Gieber, W., & Johnson, W. (1961). The city hall "beat": A study of reporter and source roles. *Journalism Quarterly, 38*(3), 289-297.

Harless, J. D. (1974). Mail call: A case study of a broadcast news gatekeeper. *Journalism Quarterly, 51*(1), 87-90.

Hirsch, P. M. (1977). Occupational, organizational, and institutional models in mass media research: Toward an integrated framework. In P. M. Hirsch, P. V. Miller, & F. G. Kline (Eds.), *Strategies for Communication Research* (Vol. 6, pp. 13-42). London: Sage.

Johnson, M. A. (1997). Predicting news flow from Mexico. *Journalism & Mass Communication Quarterly, 74*(2), 315-330.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: John Wiley & Sons.

Morrison, K. (1995). *Marx, Durkheim, Weber: Formations of Modern Social Thought*. London: Sage.

Pfeffer, J., & Salancik, G. R. (1978). *The External Control of Organizations: A Resource Dependence Perspective*. New York: Harper & Row.

Powers, A., & Fico, F. (1994). Influences on use of sources at large U.S. newspapers. *Newspaper Research Journal, 15*(4), 87-97.

Reynolds, P. D. (1971). *A Primer in Theory Construction*. New York: Macmillian Publishing.

Shoemaker, P. J. (1991). *Gatekeeping*. Newbury Park: Sage.

Snider, P. B. (1967). "Mr. Gates" revisited: A 1966 version of the 1949 case study. *Journalism Quarterly, 44*(3), 419-427.

Thompson, J. D. (1967). *Organizations in Action*. New York: McGraw-Hill.

Tuchman, G. (1972). Objectivity as strategic ritual: An examination of newsmen's notions of objectivity. *American Journal of Sociology, 77*(4), 660-679.

Tuchman, G. (1977). The exception proves the rule: The study of routine news practices. In P. M. Hirsch, P. V. Miller, & G. F. Kline (Eds.), *Strategies for Communication Research* (Vol. 6, pp. 43-62). London: Age.

Turow, J. (1990). Media industries, media consequences: Rethinking mass communication. In J. A. Anderson (Ed.), *Communication Yearbook* (Vol. 13, pp. 478-501). Newbury Park: Sage.

White, D. M. (1950). The "gate keeper": A case study in the selection of news. *Journalism Quarterly, 27*(4), 383-390.

Whitney, C. D. (1981). Information overload in the newsroom. *Journalism Quarterly, 58*(1), 69-76,161.

Whitney, C.D., & Becker, L.B. (1982). “Keeping the gates” for gatekeepers: The effects of wire news. *Journalism Quarterly, 59*(1), 60-65.

Wilke, J., & Rosenberger, B. (1994). Importing foreign news: A case study of the German Service of the Associated Press. *Journalism Quarterly, 71*(2), 421-432.

Democratic Realism, Neoconservativism, and the Normative Underpinnings of Political Communication Research

Erik P. Bucy
Department of Telecommunications
Indiana University, Bloomington

Paul D'Angelo
Department of Communication Studies
The College of New Jersey

Correspondence:

Erik P. Bucy
Dept. of Telecommunications
1229 E. Seventh Street
Indiana University
Bloomington, IN 47405-5501

Tel: (812) 856-5207
Fax: (812) 855-7955
E-mail: ebucy@indiana.edu

Paper presented to the Association for Education in Journalism & Mass Communication
Communication Theory & Methodology Division

Kansas City, Kansas
July 30, 2003

Democratic Realism, Neoconservativism, and the Normative Underpinnings of Political Communication Research

Abstract

This paper traces the development and diffusion of four basic, normative assumptions in the political communication literature identified by Chaffee and Hochheimer (1985), examining how they were brought into the field by Paul Lazarsfeld and his Columbia school colleagues under the guiding principles of democratic realism. We then show how these assumptions continue to operate in the literature today. In updating the normative orientations that inform political communication research, the analysis focuses on certain research agendas as exemplar efforts of a changed philosophical position, called neoconservatism, which undergirds and structures studies critical of press performance. In particular, we critically examine three research programs in political communication that adhere to the field's formative intellectual assumptions as enduring critiques of media and democracy: the videomalaise hypothesis, media intrusion theory, and the social erosion thesis. Suggestions are made for a fuller integration of communication perspectives into political communication research as well as new ways of thinking about media in relation to politics.

Keywords: Democratic realism, Neoconservatism, political communication research, videomalaise, media intrusion theory, social erosion thesis

In a thought-provoking commentary on the "disciplinary divide" between political science and communication research, Jamieson and Cappella (1996) observed that academic disciplines "see research through the biases created by their presuppositions and preferred methods" (p. 13). These biases, and the research traditions they produce, cause political scientists to focus primarily on outcomes and the social-economic judgments that shape them, while prompting communication researchers to study the messages that constitute campaigns (Jamieson & Cappella, 1996).[1] The subfield of *political communication*, which brings together these two perspectives, is premised on the idea that mass media institutions and messages play an intervening role in the political process. Whether conducted from a political science or mass communication orientation, political communication research tends to view media institutions as occupying an intermediary position between candidates who require media coverage to run for elective office and voters whose political behaviors depend, in large part, on information they receive from news. In the context of politics, media use is rarely regarded as an end in itself, that is, as an *outcome* of individual voter motivations (Pinkleton, Austin, & Fortman, 1998).

Historically, political communication research has been concerned with understanding the interconnections among media messages, political processes, and public involvement. The touchstone for exploring these linkages has been the political campaign setting (Davis, 1990). Since the field's inception with the classic voting studies conducted by Paul Lazarsfeld and colleagues (Lazarsfeld, Berelson, & Gaudet, 1944; Berelson, Lazarsfeld, & McPhee, 1954), political communication has been guided by certain normative orientations that cast media use as an independent variable that helps explain a desired political outcome, or dependent variable, such as civic participation, attitude formation, or vote choice. Chaffee and Hochheimer (1985) persuasively argued that these normative views take the form of certain intellectual assumptions, which involve how voters *should* act, how political institutions *should* operate, how theorists and researchers *should* do their work, and how communication and politics *should* interact. Operationally, these assumptions manifest themselves in research designs that: (a) treat the act of voting as a consumer decision and political communication research as the study of marketing problems; (b) assess the imperfect processes of politics and communication in contrast to an idealized system; (c) view the media/politics interface from a top-down or elite perspective; and, (d) push for broad generalizations that argue the processes involved in political communication are approximately equivalent across media messages, audiences/samples, time, and research settings (Chaffee & Hochheimer, 1985, pp. 268-269).

These intellectual assumptions, brought into the mainstream of the field by Lazarsfeld and his Columbia school colleagues, would not be accepted by many researchers today (Chaffee & Hochheimer, 1985, p. 269). Nevertheless, Chaffee and Hochheimer observed in the mid-1980s that "to a remarkable extent, most of the assumptions of the original studies continue to structure current research" (1985, p. 270). A fair measure of influence, they suggested, was the number of citations in collected volumes of political communication research. In the hefty, 732-page *Handbook of Political Communication* (Nimmo & Sanders, 1981), for example, Lazarsfeld and Berelson were cited in more than 40 instances; in almost all cases the reference was to their 1940 and 1948 election studies. A decade later, in the slimmer, 413-page *New Directions in Political Communication* (Swanson & Nimmo, 1990), there were still 13 references to *The People's Choice* (Lazarsfeld, Berelson, & Gaudet, 1944), *Voting* (Berelson, Lazarsfeld, & McPhee, 1954), and the follow-up study, *Personal Influence* (Katz & Lazarsfeld, 1955). And in a more recent special volume of *Research in Political Sociology* focusing on mass media and politics (Wasburn, 1995), the influence of the early voting studies could still be felt, although much less so, with seven references to the classic voting studies.

This essay first assesses how the assumptions identified by Chaffee and Hochheimer were brought into the field by the Columbia school research team, in particular by Lazarsfeld himself, and, second, examines how certain normative orientations continue to operate in the political communication literature today. Chaffee and Hochheimer (1985) make a compelling case that Lazarsfeld's "basic orientation [to] the psychology of marketing" (p. 270) and his fixation with attitude change and electoral outcomes (as opposed to information gain or candidate impression formation) as the likely outcome of media use led him to make methodological decisions that occluded a comprehensive picture of the media's role in the campaign process. In exploring the formative role of intellectual assumptions in political communication research, the influence of democratic realism on the early literature becomes evident. Specifically, we argue that the Columbia studies, via the politics-centric model of analysis they embraced, relied on the prevailing assumptions of democratic realists, who questioned the rationality of ordinary citizens and propounded the view that politics, particularly in the campaign setting, should favor the needs of politicians and parties (see Chaffee & Hochheimer, 1985). In short, the classic voting studies promulgated what we identify as a *bias of political science*, one that favored informed citizenship and fair media coverage but only insofar as each contributed to the needs of the political system.

The second part of this essay considers the enduring influence of these assumptions. Chaffee and Hochheimer were particularly sharp in their criticism that both the Columbia studies and the early versions (c. 1948-1972) of the ensuing American National Election Studies (NES) had unjustifiably relegated mass media to a minor role.[2] By limiting the number of media variables on the early NES instruments–the major vehicle of political data collection–mainstream political research inadvertently perpetuated the misguided limited-effects model of mass communication. Chaffee and Hochheimer (1985) suggested that these assumptions were resilient even though political communication research was steadily chipping away at their central propositions. For example, they noted that research continued to clarify the kinds of media use that enhance political participation and the kinds that depress it. However, they argued that those efforts were steeped in a normative perspective that assumed an idealized standard of media use "toward which the young person is expected to become socialized" (p. 292)–a standard, incidentally, that was particularly critical of television news.

In updating how these normative assumptions guide political communication research, the approach taken here is to elaborate upon certain research agendas as exemplar efforts of a philosophical position, called neoconservatism, which undergirds and structures those assumptions. In particular, we focus on press performance in political campaigns and critically examine research that (1) casts media coverage as an agent of social or political ennui (i.e. the videomalaise hypothesis); and (2) characterizes the press as an intrusive, interpretive, and adversarial part of political processes (i.e. media intrusion theory). In addition, we analyze work that (3) implicates the mass media in the erosion of social capital (i.e. the social-erosion thesis).

The Formative Role of Intellectual Assumptions

The early voting studies–what Chaffee and Hochheimer (1985) called the "original studies" of the communication field (see also Delia, 1987, p. 20)–laid the groundwork for the dominant paradigm of empirical mass communication research.[3] The potential influence of media messages, initially on the vote and later on other variables (see Berelson & Steiner, 1964; Gitlin, 1978), was found to be attenuated in these early studies by an individual's group commitments (e.g. partisan affiliation) as well as their interpersonal communication behaviors (see Rogers, 1994). Such findings gave rise to the "limited effects" theory of mass communication that prevailed early in the history of communication research. Shortly after the field was prematurely eulogized by Berelson (1959)–a colleague of Lazarsfeld's–mass communication research was put on a new footing via the synthesis provided by Klapper (1960), one of Lazarsfeld's former students. But the finer points of Klapper's analysis, particularly the contingent conditions that might produce strong, direct media effects, "tended to get rounded off in summary treatments by other authors, however, and relatively few academics saw much promise in the study of the political impact of the mass media" (Chaffee & Hochheimer, 1985, p. 288).

Chaffee and Hochheimer (1985) offer evidence that the field was shaped in its nascent period by the Columbia team's adherence to the belief that campaign messages only have effects when they *change* an individual's evaluation of a candidate (the *marketing* assumption), and the belief that political persuasion occurs in a standardized fashion across time and is homogeneous across different research settings (the *homogeneity* assumption). A chief influence Lazarsfeld seems to have had on the field of mass communication research in its formative stages was the imposition of a marketing perspective on questions of public opinion. Lazarsfeld's personal outlook took shape in his early career at the University of Vienna, where he was interested in applying survey methodology and statistical analysis to marketing problems (Chaffee &

Hochheimer, 1985, p. 270; Czitrom, 1982, Ch. 5). After Lazarsfeld relocated to the U.S. in 1935, he established the Bureau of Applied Social Research at Columbia University and, according to Katz (1987), busied himself "with a series of panel studies on the role of mass communication in the making of decisions, whether to make a vote choice, purchasing decision, or change an opinion" (pp. 525-526). Lazarsfeld's interest in decision-making was a direct continuation of his early interest in occupational choice (Katz, 1987).

In addition to advocating a marketing approach, Columbia school researchers found attractive the "growing appetite within social science" (Chaffee & Hochheimer, 1985, p. 269) to formulate broad theoretical statements, or laws, of human behavior drawn from generalizations about variables under study. "Their conclusions were always phrased in terms of general variables, in a fashion extending well beyond the empirical boundaries of their study" (Chaffee & Hochheimer, 1985, p. 277). The homogeneity assumption contributed to this tendency. In particular, the key finding that media usage primarily reinforced voting intentions was arrived at by assuming similarity across study participants. Treating each respondent as an equally weighted unit without varying susceptibility to persuasion allowed the Columbia team to interpret the low frequency of vote change over time as indicative of low levels of voter conversion and high levels of reinforcement. "It was this interpretation of marginals, coupled with the finding that those classified as 'conversion' cases were low in attention to the campaign via the media, that led to the 'law of minimal consequences' regarding media effects" (Chaffee & Hochheimer, 1985, p. 279).

Adherence to both the marketing and homogeneity assumptions led Lazarsfeld and colleagues to utilize research methods geared toward conservative measurement of media effects and state their findings in terms of general laws–even when the empirical evidence indicated more modest conclusions about the potential of interpersonal and group processes to mitigate the media's power to influence the vote (Chaffee & Hochheimer, 1985). In addition, the Columbia team worked under the guiding light of a *rationality* assumption, which stipulated that citizens should be concerned, cognizant, rational, and accepting of the political system; and, an *institutional* assumption, wherein the needs of elite groups and individuals, particularly parties and politicians, were considered paramount over the needs (informational and political) of individual voters (Chaffee & Hochheimer, 1985, p. 268). Attendant with these assumptions is the view that the news media play a democratically important supporting role and ought to be "comprehensive, accurate, and scrupulously fair and politically balanced" (Chaffee & Hochheimer, 1985, p. 268)–presumably to satisfy citizens' information needs by accurately conveying messages from politicians and political parties during election campaigns.

The Bias of Political Science

In a recent analysis of normative critiques of news and democratic processes, Bucy and D'Angelo (1999) provide perspective on how the constellation of assumptions concerning voter rationality and the proper institutional role of the press guided early communication research. They argue that the early voting studies, though empirical in orientation, were situated within an intellectual climate rife with unarticulated normative theories of the press and democracy. In particular they note the influence of the Hutchins' Commission inquiry into freedom of the press (see Hutchins, 1947), elaborated by Peterson (1956) as the social responsibility theory of the press. Although this normative theory of media and democracy crystallized in the middle of the twentieth century and was convincing for its time, political communication research began to move in another direction, namely, "along empirical, positivist, and behavioral lines of inquiry that characterized its social scientific origins in sociology, psychology, and political science" (Bucy & D'Angelo, 1999, p. 310; see also Davis, 1990).

Precipitated by the divergent impulses of the Columbia school's influential early studies–a tendency to make normative statements within the "value free" rubric of objective empirical research (see Gitlin, 1978)–an internal crisis of political communication research ensued, marked by a schism between critical and social scientific views of media and democracy (Blumler, 1983). As Bucy and D'Angelo (1999, p. 312) noted, "in the critical view political communication research itself is an expression of normative assumptions, even when it professes merely to describe or explain objective reality" (see also Lanigan & Strobl, 1981). To a large degree, this crisis was marked by the susceptibility of political communication research to the *bias of political science*–the social-scientific tendency to combine empirical methods and inductive theory-building with the assumptions of democratic realism (see Ricci, 1984; Westbrook, 1991). This biased view of democracy, which exalted political institutions while minimizing the role of citizens and mass media in political processes, examines politics "in terms of the needs of the political system, in particular the electoral component of that system, and from the perspective of political elites," in other words, from the top or power center rather than from the bottom or periphery of the system (Chaffee & Hochheimer, 1985, p. 269).

Suppressed in this intellectual climate was the development of a sophisticated theory about how citizens, political elites, and the press (as an important social institution) should interact with some sense of *mutual* dependency, a project that would only be taken up with the documentation of strong media effects and a revised view of the role relationships between political and media institutions decades later (see Blumler & Gurevitch, 1981, 1995; McCombs & Shaw, 1972). Instead, early in the field's history media institutions were generally regarded as subordinate to political institutions and, much like school, the family, and the church, were expected to lend legitimacy to democratic processes.

Democratic Realism

Democratic realism took hold as a theory of political functioning in the years following World War I (Schudson, 1983; Westbrook, 1991). The democratic realists of the 1920s, including journalist Walter Lippmann and social scientist Harold Lasswell, "focused their criticism of democracy on two of its essential beliefs: the belief in the capacity of all men for rational political action and the belief in the practicality and desirability of maximizing the participation of all citizens in public life" (Westbrook, 1991, p. 282). Here, we see a reaction against the progressive ideals of mass education and civic involvement popular with turn-of-the-century reformers (Schudson, 1998). In the realist view, democratic order is maintained via the countervailing interests and values of social organizations, interest groups, and bureaucracies, steered from above by the political realm and managed administratively by technocratic experts (see Lippmann, 1922/1965).

Stifled in this position is the need for broad citizen participation in government–and even in interest group politics.

Slowly, democratic realism won over political scientists, many who had hitherto shared the conviction of John Dewey, a vocal critic of democratic realism, that the progressive potential of social science could serve the cause of reform and couple "the interests of the hand worker with the brain worker" (Westbrook, 1991, p. 276). The mechanism for this disciplinary conversion lie in the emerging consensus within political science that individuals made important civic decisions, such as vote choice, on the basis of political attitudes that were composed of "irrational" or emotionally driven preferences and desires (Westbrook, 1991). The bias of political science, therefore, explained lapses in voter behavior, such as infrequent or shallow deliberation or inconsistent use of the news media for information, as a psychological artifact that could be corrected, presumably, by additional information or improved civic habits–and greater coordination by political elites. The zeal with which political thinkers during the 1920s and 1930s cast political processes as best managed by elites led to work, however, that served "to undermine the very object which the discipline was professionally committed to support, namely, the democratic polity" (Ricci, 1984, p. 78).

In light of these early tendencies, a clearer picture emerges as to why and how the intellectual assumptions discussed by Chaffee and Hochheimer (1985) were imported into mass communication research. As Peters (1989) points out, democratic realism–and its later expression as the theory of elite pluralism (Key, 1961)–guided communication research down an empirical path convergent with Lippmann's belief that "debates about democracy should be carried on only via the discipline of objective science" (p. 208). In this light, Lazarsfeld represents a "classic example of a transitional figure in theory development" (Baran & Davis, 2003, p. 15) in the sense that the classic voting studies advanced a "new" view of media influence: that news had but limited influence on voter decision-making. Not inconsequentially, this received view was established within an intellectual climate that championed the needs and interests of political elites and institutions over the needs and interests of individual voters and participation by the public.

Columbia school researchers did not overtly rely on the philosophical tenets of democratic realism to interpret their findings but embraced them implicitly rather than explicitly (Bucy & D'Angelo, 1999; McLeod, Kosicki, & McLeod, 1994).

The limited effects and minimal political role ascribed to the media in the early voting studies bequeathed to the field a view of the press as being chronically at odds with democratic theory. Indeed, the Columbia school studies did not shy away from suggesting that the press disserved democratic ideals. In finding that news primarily activated voters' latent predispositions toward a political party, rather than directly influencing vote choice, *The People's Choice* exploded the notion that elections were a rational decision-making process based on democratic norms of informed citizenship and civic deliberation. But the reason for underplaying the political role of the media may have less to do with what the Columbia school studies "found" than with the fact that the prevailing disciplinary outlook contained an interesting ambivalence regarding the mass media.

On one hand, the bias of political science research, with its insistence on "value-free" empirical methods, did not allow for linkages with existing normative theories of the press and democracy. If it had, the early voting studies could have steered subsequent communication research toward solving some of the problems associated with the media in politics, for example, suggesting ways that press coverage could be improved so as to help voters become more informed and invested in civic life. On the other hand, the bias of political science abided the view that the news media were primarily agents of political persuasion that should influence vote choice (i.e. the marketing assumption). Had the Columbia school studies reflected on the important *information* role of the news media and their contribution to political knowledge, as opposed to political persuasion, they might have suggested ways to amend media performance. In any case, we are still, as Schudson (1995) points out, "a long way from a coherent normative theory of journalism" (p. 29), and part of the reason why may be traced to the negligible role allowed media in mid-century conceptualizations of democratic practice, a bias of political science cultivated in the formative stages of mass communication research.

Intellectual Assumptions in Contemporary Political Communication Research

When the "intellectual leadership in U.S. election research shifted to the University of Michigan in the 1950s," Chaffee and Hochheimer (1985, p. 282) noted, the intellectual assumptions that formed the interpretive basis of the seminal voting studies went there as well. If the early classics of political research found that media had limited effects, then the door was left open for party identification and political ideology to become the core concepts in predicting the vote (see Lau & Sears, 1986). Stymied by the prevailing view of media impotence, it would take another two decades for mass communication researchers to finally cast doubt and close ranks on the limited effects model of media influence (Gitlin, 1978; McCombs & Shaw, 1972). Once again, election research was influential in redirecting the debate over media effects.

Setting out to test earlier claims that media were "stunningly successfully" in telling people what to think *about* (Cohen, 1963, p. 13), McCombs and Shaw (1972) published the first cogent statement of the agenda-setting function of mass media with their study of the 1968 presidential campaign. Agenda-setting validated widely held but empirically unsupported views dating to Lippmann's (1922/1965) time that audience awareness of and learning about political issues depends a great deal on which issues are covered and made salient in news coverage (Lang & Lang, 1959). With the rise of agenda-setting research, it became increasingly apparent that the weak effects model clashed with how media actually operated in society. Since the notion of media impotence contradicted lived experience–as well as political and journalistic folklore (Graber, 1987)–a new generation of political communication researchers redoubled efforts to demonstrate media influence on political life (e.g., Chaffee, Ward, & Tipton, 1970; McCombs & Shaw, 1972; Mendelsohn & O'Keefe, 1976; Patterson & McClure, 1976).

This renewed research effort, memorialized by Chaffee (1975) in an important volume of essays on political communication, asserted a more active role for the press in the political process and established a new base of research findings showing that media can have important cognitive and electoral influences. Arguably, the successes of empirical researchers studying political communication in the 1970s catalyzed a paradigm shift in communication research, which itself energized the "ferment in the field" (early 1980s) era of media study. During this debate, cultural scholars, perceiving a tyranny of numbers and blind faith in statistical methods, argued for a more inclusive disciplinary landscape (see Gerbner, 1983). Throughout this definitional decade, studies in political communication bolstered the case that news messages, institutions, and, increasingly, journalists themselves were central to the conduct and outcome of elections and constituted a dominant influence on the public's perceptions of candidates and issues (Graber, 1987). By the mid-1980s, Chaffee and Hochheimer (1985) observed a "new flowering of research exploring the role of mass communication in political processes, breaking out of the mold established by Lazarsfeld in a number of ways" (pp. 269-270). The case for strong media effects began to solidify: the limited effects model had become outmoded.

Over the past 20 years, research on political framing, priming, and information processing has convincingly demonstrated cognitive and emotional effects of media exposure, most effectively in experimental settings (Graber, 2001; Iyenger & Kinder, 1987; Kraus & Perloff, 1985). In a substantial reexamination of the "myth of massive media impact," Zaller (1996) has argued for a complete break with the old minimal effects tradition. "At least in the domain of political communication," he writes, "the true magnitude of the persuasive effect of mass communication is closer to 'massive' than to 'small to negligible' and...the frequency of such effects is 'often'" (Zaller, 1996, p. 18). Somewhat predictably, the shift to strong effects models of media influence has bolstered the case that media institutions need to be reigned in by the political system. No longer viewed as subordinate, news media are now seen as political actors in their own right, that is, as rival or miscast institutions (e.g. Cook, 1998; Patterson, 1993, 2002; Sparrow, 1999).

Although the received wisdom about media effects has evolved, the intellectual assumptions brought into political communication research by the early voting studies continue to shape research agendas and priorities. Just as democratic realism supported the importation of these assumptions in the field's early stages, today a different theory of democracy, *neoconservatism*, supports their continued operation. The next section traces the rise of neoconservativism as a school of political thought and describes three programs of political communication research that adhere to the field's formative intellectual assumptions as enduring critiques of media and democracy: the videomalaise hypothesis, media intrusion theory, and the social erosion thesis.

The Neoconservative Critique of Mass Media

The neoconservative critique of mass media holds that "the American press has become a permanent opposition, disparaging governmental authority, criticizing the functioning of a market economy, and producing political apathy and cynicism among the public" (Carragee, 1993, pp. 339-340). Rising levels of distrust toward politicians, and a disgust of politics in general, are directly attributed to the anti-institutional themes and relentlessly negative portrayal of political elites by the press. For neoconservative critics, the line of demarcation between a fair and balanced press and a biased, openly antagonistic news media is the Vietnam War period (roughly, the late 1960s), when journalists began to venture beyond official sources of information and started to become more interpretive in their approach to the news (Hallin, 1985).

This period of time coincides with the rise of neoconservativism as a political perspective. In 1970, Dorrien (1993) notes, the editors of *Dissent* magazine began to actively look for a term to describe "an assortment of former liberals and leftists who had recently moved to the Right" (p. 1). Regardless of their exact location on the political spectrum (e.g. the right wing of the Left or the left wing of the Right), neoconservatives were united in their disillusionment with the Johnson administration's War on Poverty and Great Society social programs–not for their intended effect of helping the have-nots and creating a more egalitarian society but for encouraging the formation of "a 'New Class' of parasitic bureaucrats and social workers" (Dorrien, 1993, p. 1).

Whereas traditional conservatives favor the outright elimination of the welfare state and a return to classical liberal conceptions of unfettered free enterprise (and the attainment of status and power through the ownership of private property and the accumulation of capital), neoconservatives prefer a minimal welfare state and seek to increase their influence primarily through organizational position (Dorrien, 1993). Hence, the substantial presence of neoconservative perspectives in policy debates, opinion journals, think tanks, and the like. As a critique of the role of intellectuals in modern society, the neoconservative conception of the New Class extends arguments developed by Schumpeter (1942) and Hayek (1949) during the New Deal era of American politics (Dorrien, 1993). Over time, the neoconservative movement has pursued a twofold political-economic and cultural agenda, which Habermas (1989) has identified as opposing communism (or disparaging socialism in favor of capitalism) and supporting the republican theory of democratic rule by traditional elites (with traditional values). Attacks on the "liberal media" are a common theme throughout much neoconservative criticism, which accuses the news media–television in particular–of possessing a political ideology that is deeply critical of political and economic authority (Lichter, Rothman, & Lichter, 1986).

Neoconservative critiques of television gained currency in the aftermath of the Vietnam War. Braestrup (1977) contends that negative coverage of the Tet Offensive in Vietnam transformed a military victory into a troubling psychological defeat. Critical reports of the war, so this argument goes, eroded public support for American foreign policy and contributed to the sense of defeat at home (Rothman, 1979). The influence of this argument can be seen in subsequent American military interventions, which have been characterized by a high degree of media management (Kellner, 1992). Robinson (1981) directly locates the problem of America's crisis of confidence during the post-Vietnam War/post-Watergate era with network news. "Our doubts about ourselves and hostility toward our institutions would be far less severe were it not for the images we receive from electronic media, more specifically, from network journalism" (Robinson, 1981, p. 314).

The Videomalaise Hypothesis

In the wake of the political and social turmoil surrounding the Vietnam War and Watergate, Robinson (1976) advanced the hypothesis that television's critical coverage of politics and government undermined trust in politicians and governmental institutions, while heightening feelings of powerlessness. Robinson based his conclusions on survey evidence that showed a positive association between television news reliance and feelings of political distrust, cynicism, and powerlessness. More specifically, he suggested six interrelated factors that he thought explained the growth of videomalaise: (1) the size of the television news audience; (2) the then-high credibility of network news; (3) the interpretive character of network news coverage; (4) the negative tone of network news reports; (5) the emphasis on conflict and violence in network coverage; and, (6) the anti-institutional themes in network news (Robinson, 1976, p. 426). Because television reached what he called a large *inadvertent audience* consisting of viewers "who fall into the news" rather than actively seek it out, and because heavy television viewing correlated with lower education and socioeconomic status, he asserted that the inadvertent audience was vulnerable to influence. "Television journalism reaches an audience more susceptible than any other," Robinson (1976, p. 430) asserted, and was comprised of "millions who would ordinarily go without the information and who lack the sophistication to deal with this journalistic content."

Following Robinson's lead, empirical political communication research began to investigate whether rising levels of political apathy and disaffection, along with declining levels of trust in government, could be directly connected to media use, particularly television viewing. Videomalaise research sparked new and sustained interest on the part of political scientists regarding the inescapable role that mass media play in election campaigns. Until the 1970s, Atkin (1981) observed, most political scientists "did not even consider the mass media as potential agents of political socialization; the few studies which included media variables treated political exposure as a criterion of socialization rather than an independent influence" (p. 301). As recently as the late 1980s, writers such as Dahl (1989) and Sartori (1987) could devote entire books to political theory without directing *any* attention to the relationship between democracy, public opinion, and mass communication (see Zolo, 1992). Collections of classic readings in American politics (e.g. Nivola & Rosenbloom, 1986) likewise overlooked the role of mass media in political life.

By the early 1990s this disciplinary outlook had changed. Reluctantly, and somewhat belatedly, mainstream political science arrived at the media table,[4] a development signified by the inclusion of a chapter devoted to political communication in the second "state of the discipline" volume published by the American Political Science Association (Graber, 1993) and by subsequent works on the news media's central (often perceived as corrosive) role in democratic processes (e.g. Cook, 1998; Graber, 2002, 2003; Patterson, 1993, 2002; Sparrow, 1999). Notwithstanding the corrosive view of the press in politics that the videomalaise hypothesis advanced, the acknowledgement of media influence was a major concession for political science to make.

Political communication research continues to investigate whether rising levels of political apathy, disaffection, and governmental mistrust can be directly connected to media use, demonstrating the continuing influence of videomalaise in shaping research agendas. However, recent research tends to qualify the theory's core claims. Rather than finding a direct link between TV news exposure and cynicism, for example, some studies have found instead that *perceptions* of media are linked with perceptions of government and that "their fortunes have declined together" (Bennett, Rhine, Flickinger, & Bennett, 1999; see also Bowen, Stamm, & Clark, 2000). Research on "political disaffection" has found that cynicism is not a uniformly negative voter attribute and may relate positively to voting intent (e.g. Austin & Pinkleton, 1995; Pinkleton & Austin, 2002). Other work has decoupled television from direct influence on viewer cynicism (Leshner & McKean, 1997) and has documented an enduring positive association between news viewing and political knowledge (Weaver, 1996; Weaver & Drew, 2001). In a wide-ranging comparative analysis of political attitudes in the U.S. and western Europe, Norris (2000) concluded that survey evidence "consistently fails to support the claim that attention to the news media in general, and television news in particular, contributes to deep-rooted indicators of civic malaise and erosion of diffuse support for the political system" (p. 314).

Media Intrusion Theory

Media intrusion arguments are characterized by two central concerns: (1) the press has assumed an inflated structural role in political processes (particularly elections) that it ought not have; and, (2) interpretive journalism, because it is not viewed as "comprehensive, accurate, and scrupulously fair and politically balanced" (Chaffee & Hochheimer, 1985, p. 268), represents a threat to the healthy functioning of the democratic system. In contrast to an idealized model of journalism, existing media practices are regarded as increasingly disruptive and an inappropriately arbitrative part of the campaign process, especially in presidential elections, where journalism has filled a vacuum created by the decline of the political parties as the primary mediating institution between politicians and the public (Kerbel, 1995; Patterson, 1993). With the introduction of direct-vote primaries to select presidential nominees after the 1968 election, political candidates were forced to appeal to heterogeneous and widely dispersed statewide audiences and became dependent upon mass media to reach voters (Patterson, 1980). This situation amplified the press' role in elections and has allowed political correspondents to act as a kind of screening committee or filtering mechanism for presidential aspirants (Davis, 1990; Schudson, 1983). Media institutions are thus held to be in competitive opposition to the political parties (Davis, 1990).

Davis (1990) has labeled this critique of journalism *media intrusion theory* and notes that it draws on theories of institutionalism developed by political scientists. In this analysis, media are evaluated as *social* institutions that should be expected to support *political* institutions, especially the parties, by allowing candidates to base campaigns on issues rather than press-defined priorities (Davis, 1990). Journalists instead, following professional norms and practices that value coverage of individuals over institutions, frame the campaign in competitive or personal terms and "devote considerable space to discussion of campaign strategy and to human interest coverage of the private lives and character of the candidates" (Davis, 1990, p. 161). Broadcast journalism, in this view, is especially structured to inform voters about compelling stories and charismatic personalities rather than policy issues or candidate positions (Chaffee, Zhao, & Leshner, 1994). Ranney (1983) maintains that the shift to television as the dominant medium of political communication may itself be the primary explanation for the decline in the salience and influence of the parties.[5]

In its extreme form, media intrusion theory maintains that news media have become political actors in their own right (e.g. Cook, 1998; Sparrow, 1999) and that the news media compound the problem by explicitly covering press-candidate interactions via cynical, self-reflexive narratives. Such narratives are believed to constitute a "strategic haze" (Kerbel, 1997) that drowns out coverage of issue stories, in the process cultivating views among news consumers that candidates cannot be trusted and that the press is oriented to see politics merely as a game with winners and losers rather than a civic process though which candidates address and seek to resolve problems (see also Kerbel, Apee, & Ross, 2000).

When political scientists discuss the substance of media coverage, they tend to be quite critical of journalism's performance, assessing news coverage in view of the political system's needs rather than the attempts by journalists to maintain a certain degree of editorial autonomy. As Graber (1987) observes, "they have complained about the heavy emphasis on 'horse race and hoopla,' the de-emphasis of issues, and the large number of stories dwelling on the personal qualities of the candidates" (p. 12). Content analyses of campaign stories carried out by political scientists have found that less than a third of campaign coverage even mentions issues (Kerbel, Apee, & Ross, 2000; Patterson, 1980; Robinson & Sheehan, 1983), compared to a preponderance of political strategy coverage. Patterson (1993, 2002), moreover, has documented a negative and evaluative pattern of media coverage of politics. These findings, and the coding strategies that produced them, point to a disruptive role for political journalism and suggest the news media engage primarily in an adversarial relationship with political power (Carragee, 1993; Davis, 1990). But they discount the increasing amount of political orchestration visited upon reporters by political handlers and image consultants (see Blumler & Gurevitch, 1995; Sabato, 1981).

The adversarial position resonates with a wide audience (not the least of which are journalists) because to some extent both the press and political actors view themselves in these terms (Blumler & Gurevitch, 1981). The metaphors of the press as the "Fourth Estate" or "watch dog" on government stem from this professional ethos or ideology. Champion of the public's right to know, the adversarial press sees itself as an independent check on the political system, a seeker after truth that ferrets out evidence of official corruption and ineptitude and acts as a guardian against tyranny. However, there are several structural limitations to the adversarial model. First, as Blumler and Gurevitch (1981) point out, it accommodates just one mode of interaction between media and politicians: antagonism. Secondly, the adversarial explanation ignores the mutual dependency that journalists share with political actors; it fails, as Grossman and Rourke (1976) observe, to provide a "mechanism for understanding the enormous amount of cooperation and even collaboration that takes place in the interaction between the press... and the government" (p. 455). Finally, if political message-making is a joint enterprise, a strict adversarial stance could not be sustained for any length of time without eroding the very basis of the relationship (Blumler & Gurevitch, 1981).

Still, concern over the media's role in politics as potentially harmful to democracy underscores an important aspect of the bias of political science in its contemporary neoconservative form; namely, that media institutions should be subordinate to political institutions, not rivals for the public's attention. More contextually, media intrusion criticisms evoke Lippmann's (1922/1965) ideas on the limits of ordinary citizens to understand complex events and his belief that journalism could best serve society by supplying experts with information needed to make intelligent governing decisions. Media intrusion theory thus complements the elite pluralist view of democracy, which stems in part from Lippmann's writings.

Recent communication research continues to validate the allure of intrusionist arguments. In *The Press Effect*, Jamieson and Waldman (2003) argue that the broadcast news networks, as a shaper of political events, subverted the political process by miscalling the 2000 presidential election. On election night, they maintain, the eagerness to call a winner–first Al Gore and then George W. Bush–not only misled the public but also may have shaped the post-election battle. The election's outcome "was by no means the forgone conclusion that some in the press would have had us believe. Had they happened differently, any number of events and decisions could have resulted in a Gore victory" (Jamieson & Waldman, 2003, p. 76). For another example of the media intrusion perspective at work we do not have to look farther than some of Chaffee's own research. In an analysis of television talk shows from the 1992 presidential election, Chaffee, Zhao, and Leshner (1994) suggest that these media formats contribute to the gradual erosion of party authority. "The undermining of political parties as electoral institutions is a long-term effect of television feared by thoughtful political scientists (e.g. Ranney, 1983), and the extensive interview shows of 1992 seem to have done nothing to reverse this trend" (Chaffee, Zhao, & Leshner, 1994, 318). The assumptions of political science continue to have influence in shaping research at the media/politics interface, even when conducted from a communication perspective.

Thomas Patterson, who has built a career assailing the press' role in four influential, and suggestively titled, books–*The Unseeing Eye* (1976, with McClure), *The Mass Media Election* (1980), *Out of Order* (1993), and *The Vanishing Voter* (2002)–is perhaps the leading critic of media and politics writing in the media intrusion tradition today. Patterson regards the press as a jaded, miscast institution, one that is neither democratically accountable nor very well suited for coalition building–a major function of elections. Commenting on the news media's arbitrative character, he observes:

> The media have a powerful drive toward skepticism, a persistent need for novelty, and a weakness for personality. These are deeply rooted and longstanding tendencies within the media that are based on their need to attract an audience and to avoid taking sides in partisan conflict. These tendencies, however, are not the desired characteristics of an electoral intermediary. What voters need from an intermediary is enough consistency and permanence to enable them to keep their eyes on the horizon. What they get instead from the press is a version of politics that centers on incidents and interruptions (2002, p. 62).

Similar to other neoconservative media critics, Patterson argues that the problem of the modern presidential campaign lies primarily in the role assigned to the press and not with other important players in the process, such as political action committees or the political consulting or advertising industries. The press, he says, is ill-suited for the role of democratic broker and imposes its own values on American politics.

In this view, journalistic values are at odds with political values, which results in a news agenda that misrepresents what is at stake (Patterson, 1993, 2002). They also introduce an element of "random partisanship" (or personality politics) into campaigns. Moreover, election news drives a wedge between candidates and voters rather than serving to bring them together. Hence, political journalism as currently practiced violates the assumption that media coverage should be comprehensive, scrupulously fair, and politically balanced. Other writers have not been so circumspect. Rothman, for instance, has written that the national news media's political role is not only inappropriate, it has directly "contributed to the decay of traditional political and social institutions" (1979, p. 346).

The Social Erosion Thesis

The position that news coverage leads to voter cynicism and alienation, contributing to the erosion of public confidence in political institutions, is fundamental to neoconservative critiques of media and society. When combined with the tendency to generalize and make broad conclusions, a feature of the early voting studies identified by Chaffee and Hochheimer, social erosion arguments become a sweeping indictment of media in society. Social erosion arguments are expansive in the sense that champions of this position are not confined to political science. They appear in slightly different form in the field of communication wherever an assumption of strong media effects and a robust normative view of society (i.e. what makes for a "strong democracy") intersect. Communication titles such as *Spiral of Cynicism: The Press and the Public Good* (Cappella & Jamieson, 1997), *Seducing America: How Television Charms the Modern Voter* (Hart, 1999), and *Media Unlimited: How the Torrent of Images and Sounds Overwhelms Our Lives* (Gitlin, 2002) exemplify this genre. What each has in common is the attempt to associate media exposure with social and political disaffection.

This position was in full evidence in the initial accounts of Harvard political scientist Robert Putnam's "bowling alone" thesis–that mass media have substantially contributed to declining levels of citizen engagement in civic affairs. Putnam (1995) asserted that the introduction of television into American society in the 1950s was a major factor in the subsequent decline of social trust, as well as membership and participation in civic organizations (i.e. social capital). Across educational level, Putnam (1995) found a negative association between the amount of overall television exposure and the level of reported social trust and number of groups an individual joins. Thus, he argued that the country's supply of social capital had eroded. With the more recent publication of his book, *Bowling Alone* (Putnam, 2000), he now distinguishes between news and entertainment television, crediting news exposure with increased civic involvement. Though Putnam no longer stands by his earlier blanket indictment of television, his position typifies the bias of political science in studies of media and democracy–in one way or another, the media are to blame for the perceived decline in civic involvement. One does not have to look far or very closely to see a strong normative orientation at work.

Putnam's argument points to two of the assumptions underlying political communication research identified by Chaffee and Hochheimer (1985), namely, that people should be concerned and accepting of the political system and that the role of media should be conceived in terms of what they might do *to* people rather than what people might be doing *with* media. In Putnam's (1995) assessment, media use is thought to displace social activities outside the home, "especially social gatherings and informal conversations" (p. 679). The notion that media use might qualify as civic engagement is normatively inconsistent with this view, although Putnam (2000) now concedes that, "as a technical matter, the extraordinary power of television can encourage as well as discourage civic involvement" (p. 410). Both in Putnam's earlier work and *Bowling Alone*, newspaper reading receives no such criticism. Instead, time spent with newspapers is associated with high levels of social capital, as it is positively related to social trust and group membership. In Putnam's reformulation of the bowling alone thesis, newspapers occupy a privileged position over television and public affairs coverage is held in higher esteem than entertainment programming, betraying an elitist view about what constitutes valid information. In the conclusion of *Bowling Alone*, Putnam expresses particular concern about electronic media:

> No sector of American society will have more influence on the future state of our social capital than the electronic mass media and especially the Internet. If we are to reverse the adverse trends of the last three decades in any fundamental way, the electronic entertainment and telecommunications industry must become a big part of the solution instead of a big part of the problem (p. 410).

As several scholars have observed (Bucy, D'Angelo, & Newhagen, 1999; Muir, 1992; Schudson, 1996), the problem of declining civic participation may in part be methodological; that is, participation depends to a large degree on the way criterion variables are selected and defined. Political science has traditionally defined participation primarily as active outdoor behaviors pursued in public rather than in private. Verba and Nie (1987), for example, identify four major participation variables: voting, campaign participation, community activities, and leader/legislator contact. Media use is not regarded as a major avenue of involvement. Similarly, Verba, Schlozman, and Brady (1995) do not consider media use in their comprehensive study of civic participation.

By defining democratic legitimacy almost strictly in terms of outdoor behaviors, chief of which is voting (Chaffee & Hochheimer, 1985), political science is confronted with the problem of either revising its measures or clinging to outdated notions of popular consent (see Bucy & Gregson, 2001)–a problem more easily identified than solved, given the reliance on standardized questions used in much survey research. Arguments for measurement stasis point out that if the question wording of the political participation items on the National Election Studies battery were to change, for example, trend analysis would no longer be possible. In this vein, the early "classics" continue to exert considerable influence on the conceptualization and execution of much political communication research. Despite a growing awareness of media centrality, assumptions about the preferred marginal role of the press in political life persist. And historical disciplinary requirements, stemming from the normative orientation of analyzing politics in terms of the needs of the political system rather than individual voters, continue to hold sway. Traditional conceptions of political participation, however, may not go far enough in explaining actual citizen involvement in democratic processes.

Rather than "relegating media-related activity to the status of a minor mode of political participation," as much political research has done for decades through the National Election Studies (Chaffee & Hochheimer, 1985, p. 284), media involvement might instead be treated as a primary or major mode of civic participation, that is, as a political behavior in its own right. The process of civicly engaging audiences through new media use has been labeled "media participation" (Bucy & Gregson, 2001). In addition to providing access to news and information, the new media, especially call-in programs and civically oriented sites on the World Wide Web, provide a public space for citizens to debate issues and express their support for, or discontent with, policies or office holders. The Internet also provides citizens with a ready avenue of access to decision makers, even if they choose not to reply. Regardless of whether new media use has a direct impact on political decisions and policy making, the media audience may still benefit from the continuous opportunities for civic engagement that new media provide. By enabling both horizontal communication among citizens and vertical communication with elites, interactive media perform the valuable service of giving citizens the opportunity to act, first and foremost, *as* citizens. The public sphere thus benefits from opportunities for civic activity through media, even though citizen involvement by traditional standards remains indirect (Bucy & Gregson, 2001).

Conclusion

Despite the interdisciplinary nature and diversification of political communication research into such areas as media framing and priming, agenda setting, information processing, reception analysis, and critical theory (Bucy & D'Angelo, 1999), the field has not entirely left behind the once (and many say still) dominant "voter persuasion paradigm" of media having effects on voting choices (Nimmo & Swanson, 1991). As this analysis has shown, research at the media/politics interface is driven by basic normative orientations, or biases, that stem largely from the disciplinary assumptions of political science but which are embraced by communication. Whether explicit or implicit, these biases frame many of the questions, and thus many of the findings, of political communication research. Moreover, more than one normative orientation may be at work in analyses of media and democracy at any one time.

As Graber (1987) suggested in these very pages,[6] political communication researchers on both sides of the disciplinary divide "need to become better acquainted with each other's work so that their combined efforts can produce superior findings in this complex and fluctuating research area" (p. 10). Although the problem of "shocking mutual ignorance or disregard" between political science and communication research that Graber observed in 1987 has subsided since that writing, research at the media/politics interface continues to be hampered by the constraints placed on research by the intellectual assumptions that constitute the bias of political science. This is not to say that normative assumptions are unnecessary; to the contrary, mature research cannot proceed without them. But a realistic perspective regarding the media/politics interface cannot be achieved unless political communication researchers more fully acknowledge that mass media are increasingly indistinguishable from, and enmeshed within, political processes themselves.

Overcoming these conceptual constraints will first involve a fuller recognition of the importance of media use in democratic processes. Indeed, as the political theorist Zolo (1992) has observed, for democratic theory to be properly retooled to suit contemporary conditions, political theory "should turn its central most attention to the political effects of mass communication" (p. 153). Iyengar and Kinder (1987, p. 3) similarly argue that the lack of a (universal) theory of media effects significantly impedes our understanding of how a mass democracy even works. With a massive increase in the availability and use of information and communication technologies (Gitlin, 2002), it has become impossible to ignore the cognitive, emotional, and behavioral effects mass media have on citizens of advanced industrial societies *as well as* the impact these effects have on the functioning of modern political systems. Whether seen as desirable or not, the electronic information and media environment now shapes and surrounds political life–and it is continually evolving and expanding. Thus, fully understanding the media/politics interface will require a more nuanced appreciation for the communication technology side of the equation than either neoconservative analyses or the voter persuasion paradigm has evidenced thus far.

The primary civic contribution of new media formats (talk radio, call-in television, townhall forums, political entertainment programs, and Web sites), for example, is that they seem to have made a professionalized and distant political system more accessible. In contrast to passive spectatorship encouraged by traditional media, citizen involvement in interactive environments presents voters with a ready avenue of participation across a variety of communication modalities–and on a continuous, everyday basis, not just during election campaigns. Indeed, rather than being slighted by a form of "pseudo participation," as traditionalists (e.g. Kerbel, 1999) might characterize new media use, the citizen benefits from the awareness that active involvement with new media "provides proximity to political elites, makes politics continuously available and entertaining (i.e. accessible), offers open-mike access to a wide audience, socializes citizens to participate in public affairs, and allow voters to cultivate a civic identity and know their own minds" (Bucy & Gregson, 2001, p. 375).

Thus, the political effects of media use may be far more consequential on an individual level, in terms of producing social, psychological, and civic rewards for citizens, than on policy or electoral outcomes. Such a view places the voter or individual citizen at the center of inquiry as an active audience member rather than assuming a unidirectional flow of news that "acts on" unsuspecting consumers of political communication, engendering cynicism and other ill outcomes.

A second step to more accurately explicating the role of the press in politics may lie in reconsidering the *interdependence* of media and political systems. In neoconservative thought, news coverage imparts a cynical frame of reference on campaign events and governing processes. Consumption of strategy-oriented news coverage, in turn, is believed to cultivate cyncism, or citizen mistrust toward and lack of confidence in political officials and institutions (Austin & Pinkleton, 1999; Goldfarb, 1991). Describing this tendency, journalist E. J. Dionne (1991) argues that Americans are subjected to a steady stream of anti-government themes and sloganeering during elections. Candidates, he charges, "have adopted a cynical stance that…plays into popular wisdom about politics and wins them votes" (p. 17). In response, the press has embraced an increasingly self-conscious style of reporting that describes journalists and political candidates as being engaged in a dance of mutual manipulation and mistrust (Kerbel, Apee, & Ross, 2000).

Research that critiques the videomalaise hypothesis and the intrusionist position has begun to rethink the point that the mutual dependency of the press and politicians leads to cynicism in news and corrodes the civic outlook of citizens. In this vein, Paletz (2002, p. 222) argues that journalists *reflect* the uncivil tenure of candidate discourse more so than add to it, observing that most content analyses are not sensitive to this important distinction. Similarly, Esser and D'Angelo (2003) found that press self-coverage is not uniformly cynical; rather, self-coverage is composed of various narrative frames which are indicative of an effort to demonstrate accountability to standards of professional practice. Moreover, cynicism does not necessarily depress one's voting intentions nor sense of efficacy. Austin and Pinkelton (1995) found, for instance, that cynicism may bolster an individual's belief in the ability to observe the dissimulations of politicians, which, in turn, enhances the desire to vote and the feeling that voting matters.

Disparate lines of research thus underscore a movement in the field to amend neoconservative views. These studies not only add up to a cogent critique of neoconservative ideas, they also refashion assumptions concerning the importance of media use as a form of political participation and news institutions as constructive civic actors. Given the normative biases at work in the political communication literature, researchers should more often consider counter arguments when discussing media effects rather than reflexively interpreting results consistent with the bias of political science. For new research vistas to present themselves, and knowledge about political communication processes and effects to advance, mass media should not automatically be regarded as subordinate to traditional political institutions but should be seen as an integral part of an interdependent political communication process. New normative conceptions, in turn, may then override neoconservative views that admonish the interpretive impulse of the news media, repudiate the constructive role of cynicism and self-coverage, and constrict the potential for political deliberation afforded by the news media.

As with the outmoded limited effects model, media's time at the margins of democratic theory and as a minor mode of participation has long since passed. Researchers would do well to use this observation as their departure point when designing studies that attempt to explain political communication processes in a media saturated society.

References

Atkin, C. K. (1981). Communication and political socialization. In D. D. Nimmo & K. R. Sanders (Eds.), *Handbook of political communication* (pp. 299-328). Beverly Hills: Sage.

Austin, E. A., & Pinkleton, B. E. (1995). Positive and negative effects of political disaffection on the less experienced voter. *Journal of Broadcasting and Electronic Media, 39*(2), 215-235.

Austin, E. A., & Pinkleton, B. E. (1999). The relation between media content evaluations and political disaffection. *Mass Communication & Society, 2*(3/4), 105-122.

Baran, S. J., & Davis, D. K. (2003). *Mass communication theory: Foundations, ferment, and future*. Belmont, CA: Wadsworth.

Bennett, S. E., Rhine, S. L., Flickinger, R. S., & Bennett, L. L. M. (1999). "Video malaise" revisited: Public trust in the media and government. *Press/Politics, 4*(4), 8-23.

Berelson, B. R. (1959). The state of communication research. *Public Opinion Quarterly, 23*, 1-6.

Berelson, B., Lazarsfeld, P. L., & McPhee, W. N. (1954). *Voting: A study of opinion formation in a presidential campaign*. Chicago: University of Chicago Press.

Berelson, B., & Steiner, G. (1964). *Human behavior: An inventory of scientific findings*. New York: Harcourt Brace and World.

Blumler, J. G. (1983). Communication and democracy: The crisis beyond and the ferment within. *Journal of Communication, 33*(3), 166-173.

Blumler, J. G., & Gurevitch, M. (1981). Politicians and the press: An essay on role relationships. In D. Nimmo & K. Sanders (Eds.), *Handbook of political communication* (pp. 467-493). Beverly Hills: Sage.

Blumler, J. G., & Gurevitch, M. (1995). *The crisis of public communication*. London: Routledge.

Bowen, L, Stamm, K., & Clark, F. (2000). Television reliance and political malaise: A contingency analysis. *Journal of Broadcasting & Electronic Media, 44*(1), 1-15.

Braestrup, P. (1977). *Big story: How the American press and television reported and interpreted the crisis of Tet 1968 in Vietnam and Washington.* Boulder: Westview Press.

Bucy, E. P., & D'Angelo, P. (1999). The crisis of political communication: Normative critiques of news and democratic processes. *Communication Yearbook 22*, 301-339.

Bucy, E. P., D'Angelo, P., & Newhagen, J. E. (1999). The engaged electorate: New media use as political participation. In L. L. Kaid & D. G. Bystrom (Eds.), *The electronic election: perspectives on 1996 campaign communication* (pp. 335-347). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bucy, E. P., & Gregson, K. (2001). Media participation: A legitimizing mechanism of mass democracy. *New Media & Society, 3*(3), 359-382.

Cappella, J. N., & Jamieson, K. H. (1997). *Spiral of cynicism: The press and the public good.* New York: Oxford University Press.

Carragee, K. M. (1993). A critical evaluation of debates examining the media hegemony thesis. *Western Journal of Communication, 57*(3), 330-348.

Chaffee, S. H. (Ed.). (1975). *Political communication: Issues and strategies for research.* Beverly Hills, CA: Sage.

Chaffee, S. H., & Hochheimer, J. L. (1985). The beginnings of political communication research in the United States: Origins of the "limited effects" model. In E. M. Rogers & F. Balle (Eds.), *The media revolution in America and Western Europe* (pp. 267-296). Norwood, NJ: Ablex.

Chaffee, S. H., Ward, S., & Tipton, L. (1970). Mass communication and political socialization. *Journalism Quarterly, 47*, 647-659.

Chaffee, S. H., Zhao, X., & Leshner, G. (1994). Political knowledge and the campaign media of 1992. *Communication Research, 21*(3), 305-324.

Cohen, B. C. (1963). *The press and foreign policy.* Princeton: Princeton University Press.

Cook, T. E. (1998). *Governing with the news: The news media as a political institution.* Chicago: University of Chicago Press.

Dahl, R. A. (1989). *Democracy and its critics*. New Haven: Yale University Press.

Davis, D. K. (1990). News and politics. In D. L. Swanson & D. Nimmo (Eds.), *New directions in political communication* (pp. 147-184). Newbury Park: Sage.

Delia, J. G. (1987). Communication research: A history. In C. R. Berger & S. H. Chaffee (Eds.), *Handbook of communication science* (pp. 20-98). Newbury Park, CA: Sage.

Dionne, E. J. (1991). *Why Americans hate politics*. New York: Simon and Schuster.

Dorrien, G. (1993). *The neoconservative mind: Politics, culture, and the war of ideology*. Philadelphia: Temple University Press.

Esser, F., & D'Angelo, P. (2003). Framing the press and the publicity process: A content analysis of meta-coverage in Campaign 2000 network news. *American Behavioral Scientist, 46*(5), 617-641.

Gerbner, G. (Ed.). (1983). Ferment in the field. [Special Issue]. *Journal of Communication, 33*(3).

Gitlin, T. (1978). Media sociology: The dominant paradigm. *Theory and Society, 6*, 205-253.

Gitlin, T. (2002). *Media unlimited: How the torrent of images and sounds overwhelms our lives*. New York: Henry Holt and Company.

Goldfarb, J. C. (1991). *The cynical society*. Chicago: University of Chicago Press.

Graber, D. A. (1987). Researching the mass media-elections interface: A political science perspective. *Mass Comm Review, 14*(1), 3-19.

Graber, D. A. (1993). Political communication: scope, progress, promise. In A. W. Finifter (Ed.), *Political science: The state of the discipline II* (pp. 305-332). Washington, DC: American Political Science Association.

Graber, D. A. (2001). *Processing politics: Learning from television in the Internet age*. Chicago: University of Chicago Press.

Graber, D. A. (2002). *Mass media and American politics*, 6th ed., Washington, DC: CQ Press.

Graber, D. A. (2003). The media and democracy: Beyond myths and stereotypes. *Annual Review of Political Science, 6*, 139-160.

Grossman, M. B., & Rourke, F. E. (1976). The media and the presidency: An exchange analysis. *Political Science Quarterly, 91*(3), 455-470.

Habermas, J. (1989). *The new conservatism: Cultural criticism and the historians' debate*. Ed. and trans. S.W. Nicholsen. Cambridge: MIT Press.

Hallin, D. (1985). The American news media: A critical theory perspective. In J. Forrester (Ed.), *Critical theory and public life* (pp. 121-146). Cambridge: MIT Press.

Hart, R. P. (1999). *Seducing America: How television charms the modern voter* (rev. ed.). Thousand Oaks, CA: Sage.

Hayek, F. A. (1949). The intellectuals and socialism. *The University of Chicago Law Review, 16*(3), 417-433.

Hutchins, R. (1947). *A free and responsible press*. Chicago: University of Chicago Press.

Iyengar, S., & Kinder, D. (1987). *News that matters*. Chicago: University of Chicago Press.

Jamieson, K. H., & Cappella, J. N. (1996). Bridging the disciplinary divide. *P.S.: Political Science & Politics, 29*(1), 13-17.

Jamieson, K. H., & Waldman, P. (2003). *The press effect: Politicians, journalists, and the stories that shape the political world*. New York: Oxford University Press.

Katz, E. (1987). Communications research since Lazarsfeld. *Public Opinion Quarterly, 51*, S25-S45 (supplement: 50th anniversary issue).

Katz, E., & Lazarsfeld, P. (1955). *Personal influence*. New York: The Free Press.

Kellner, D. (1992). *The Persian Gulf TV war*. Boulder, CO: Westview Press.

Kerbel, M. R. (1995). *Remote and controlled: Media politics in a cynical age*. Boulder: Westview Press.

Kerbel, M. R. (1997). The media: Viewing the campaign through a strategic haze. In M. Nelson (Ed.), *The elections of 1996* (pp. 81-105). Washington, DC: CQ Press.

Kerbel, M. R. (1999). *Remote & controlled: Media politics in a cynical age* (2nd ed.). Boulder, CO: Westview Press.

Kerbel, M. R., Apee, S., & Ross, M. (2000). PBS ain't so different: Public broadcasting, election frames, and democratic empowerment. *Press/Politics, 5*(4), 8-32.

Key, V. O. (1960). *Public opinion and American democracy*. New York: Knopf.

Klapper, J. (1960). *The effects of mass communication*. New York: The Free Press.

Kraus, S., & Perloff R. M. (Eds.) (1985). *Mass media and political thought: An information-processing approach*. Beverly Hills: Sage.

Lanigan, R. L., & Strobl, R. L. (1981). A critical theory approach. In D. Nimmo & K. R. Sanders (Eds.), *Handbook of political communication* (pp. 142-167). Beverly Hills, CA: Sage.

Lang, K. & Lang, G. E. (1959). The mass media and voting. In E. Burdick & A. J. Brodbeck (Eds.), *American voting behavior* (pp. 217-235). Glencoe, IL: The Free Press.

Lau, R. L., & Sears, D. O. (1986). Social cognition and political cognition: Past, present, and future. In R. R. Lau & D. O. Sears (Eds.), *Political cognition* (347-366). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lazarsfeld, P., Berelson, B., & Gaudet, H. (1944). *The people's choice*. New York: Duell, Sloan and Pearce.

Leshner, G., & McKean, M. L. (1997). Using TV news for political information during an off-year election: Effects on political knowledge and cynicism. *Journalism & Mass Communication Quarterly, 74*(1), 69-83.

Lichter, S. R., & Rothman, S., with Lichter, L. S. (1986). *The media elite: America's new power brokers*. Bethesda, MD: Adler and Adler.

Lippmann, W. (1922/1965). *Public opinion*. New York: The Free Press.

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly, 36*(2), 176-187.

McLeod, J. M., Kosicki, G. M., & McLeod, D. M. (1994). The expanding boundaries of political communication effects. In J. Bryant & D. Zillmann (Eds.), *Media effects: Advances in theory and research* (pp. 123-162). Hillsdale, NJ: Lawrence Erlbaum Associates.

McQuail, D. (1994). *McQuail's mass communication theory* (4th ed.). London: Sage.

Mendelsohn, H., & O'Keefe, G. (1976). *The people choose a president: Influences on voter decision making*. New York: Praeger.

Muir, J. K. (1992, October). *Teledemocracy and citizenship: Redefining public participation*. Paper presented at the annual meeting of the Speech Communication Association, Chicago, IL.

Nimmo, D. D., & Sanders, K. R. (1981). The emergence of political communication as a field. In D. D. Nimmo & K. R. Sanders (Eds.), *Handbook of political communication* (pp. 11-36). Beverly Hills: Sage.

Nimmo, D. D., & Swanson, D. L. (1991). The field of political communication: Beyond the voter persuasion paradigm. In D. L. Swanson & D. D. Nimmo (Eds.), *New directions in political communication* (pp. 7-47). Newbury Park, CA: Sage.

Nivola, P. S., & Rosenbloom, D. H. (1986). *Classic readings in American politics*. New York: St. Martin's Press.

Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. Cambridge: Cambridge University Press.

Paletz. D. (2002). *The media in American politics* (2nd ed.). New York: Longman.

Patterson, T. E. (1980). *The mass media election: How Americans choose their president*. New York: Praeger.

Patterson, T. E. (1993). *Out of order*. New York: Knopf.

Patterson, T. E. (2002). *The vanishing voter: Public involvement in an age of uncertainty*. New York: Alfred A. Knopf.

Patterson, T. E., & McClure, R. D. (1976). *The unseeing eye: The myth of television power in national elections*. New York: Putnam.

Peterson, T. (1956). The social responsibility theory of the press. In F. S. Siebert, T. Peterson, & W. Schramm (Eds.), *Four theories of the press* (pp. 73-103). Urbana, IL: University of Illinois Press.

Peters, J. D. (1989). Democracy and American mass communication theory: Dewey, Lippmann, Lazarsfeld. *Communication, 11*, 199-220.

Pinkleton, B. E., Austin, E. W., & Fortman, K. K. J. (1998). Relationships of media use and political disaffection to political efficacy and voting behavior. *Journal of Broadcasting & Electronic Media, 42*(1), 34-49.

Pinkleton, B. E., & Austin, E. W. (2002). Exploring relationships among media use frequency, perceived media importance, and media satisfaction in political disaffection and efficacy. *Mass Communication & Society, 5*(2), 141-163.

Pomper, G. N. (1977). The decline of partisan politics. In L. Maisel & J. Cooper (Eds.), *The impact of the electoral process* (pp. 14-). Beverly Hills: Sage.

Putnam, R. D. (1995). Tuning in, tuning out: The strange disappearance of social capital in America. *PS: Political Science & Politics, 28*(4), 664-683.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community.* New York: Simon & Schuster.

Ranney, A. (1983). *Channels of power: The impact of television on American politics.* New York: Basic Books.

Ricci, D. M. (1984). *The tragedy of political science: Politics, scholarship, and democracy.* New Haven: Yale University Press.

Robinson, M. J. (1976). Public affairs television and the growth of political malaise: The case of "The Selling of the Pentagon." *American Political Science Review, 70*, 409-432.

Robinson, M. J. (1981). Reflections on the nightly news. In R. Adler (Ed.), *Understanding television: Essays on television as a social and cultural force* (pp. 313-324). New York: Praeger.

Robinson, M. J., & Sheehan, M. A. (1983). *Over the wire and on TV: CBS and UPI in campaign '80*. New York: Russell Sage.

Rogers, E. M. (1994). *A history of communication study: A biographical approach*. New York: The Free Press.

Rothman, S. (1979). The news media in post-industrial America. In S. M. Lipset (Ed.), *The third century America as a post-industrial society* (pp. 346-388). Chicago: University of Chicago Press.

Sabato, L. J. (1981). *The rise of political consultants: New ways of winning elections*. New York: Basic Books.

Sartori, G. (1987). *The theory of democracy revisited, vol. 2*. Chatham, NJ: Chatham House Publishers.

Schudson, M. (1983). *The news media and the democratic process*. New York: Aspen Institute for Humanistic Studies.

Schudson, M. (1995). *The power of news*. Cambridge, MA: Harvard University Press.

Schudson, M. (1996). What if civic life didn't die? Controversy, unsolved mysteries: The Tocqueville files. *The American Prospect, 25*, 17-20.

Schudson, M. (1998). *The good citizen: A history of American civic life*. New York: The Free Press.

Schumpeter, J. (1942). *Capitalism, socialism, and democracy*. New York: Harper & Brothers.

Sparrow, B. H. (1999). *Uncertain guardians: The news media as a political institution*. Baltimore: The Johns Hopkins University Press.

Swanson, D. L., & Nimmo, D. (Eds.). (1990). *New directions in political communication*. Newbury Park, CA: Sage.

Verba, S., & Nie, N. H. (1987). *Participation in America: Political democracy and social equality*. Chicago: University of Chicago Press.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995) *Voice and equality: Civic volunteerism in American politics*. Cambridge, MA: Harvard University Press.

Wasburn, P. C. (Ed.). (1995). *Mass media and politics. Research in political sociology, vol. 7*. Greenwich, CT: JAI Press.

Weaver, D. H. (1996). What voters learn from media. *Annals, 546*, 34-47.

Weaver, D. H., & Drew, D. (2001). Voter learning and interest in the 2000 presidential election: Did the media matter? *Journalism & Mass Communication Quarterly, 78*(4), 787-798.

Westbrook, R. B. (1991). *John Dewey and American democracy*. Ithaca: Cornell University Press.

Zaller, J. (1996). The myth of massive media impact revived: New support for a discredited idea. In. D. C. Mutz, P. M. Sniderman, & R. A. Brody (Eds.), *Political persuasion and attitude change* (pp. 17-78). Ann Arbor: University of Michigan Press.

Zolo, D. (1992). *Democracy and complexity: A realist approach*. University Park, PA: Pennsylvania State University Press

Endnotes

[1] Graber (1987) has similarly noted that political science is strongest when assessing factors that influence voters' political attitudes and voting decisions and weakest when analyzing media content elements.

[2] One reason for the suppression of media variables was the disciplinary perspective of the original NES director. From 1948 until the late 1970s, the National Election Studies were directed by a psychologist, Angus Campbell, who "viewed the determinants of the vote as psychological in nature, and party identification was preeminent among these" (Chaffee & Hochheimer, 1985, p. 283). Not until after the 1972 presidential campaign, where the influence of media was unmistakable and widely acknowledged, did additional media-related measures get added to the NES questionnaires (Chaffee & Hochheimer, 1985, p. 287).

[3] Though often referenced loosely, the dominant paradigm of communication research specifically refers to an analytical outlook characterized by a liberal-pluralist ideal of society, a functionalist perspective on media, a linear transmission model of effects, a powerful view of media effects modified by group relations, and quantitative methods and variable analysis (McQuail, 1994, pp. 45-48; see also Gitlin, 1978).

[4] Of course, this statement is a generalization. Many political scientists, such as Doris Graber (a former journalist), C. Richard Hofstetter, Shanto Iyenger, Thomas Patterson, and the late Murray Edelman have been studying mass media in politics for decades.

[5] Although Pomper (1977), Schudson (1983) and others have persuasively argued that the decline of the political parties in the United States has varied and diverse causes, Ranney's position that the press played a leading role in the parties' demise is typical of writers in the media intrusion tradition.

[6] The journal in which Graber's article was published, *Mass Comm Review*, was the precursor to *Mass Communication and Society*.

The World Wide Web of Sports: A Path Model Examining How Online Gratifications and Reliance Predict Credibility of Online Sports Information

Thomas J. Johnson, Ph.D.
Professor
School of Journalism
Southern Illinois University
Carbondale, IL 62901
(618) 453-3279
fax: (618) 453-5200
tjohnson@siu.edu

Barbara K. Kaye, Ph.D.
Assistant Professor
Department of Broadcasting
University of Tennessee
Knoxville, TN 37996-0333
865-974-1118
bkk@utk.edu

Paper presented to the Comm Theory and Methodology Division of the Association for Education in Journalism & Mass Communication, Kansas City, Mo., August 2003.

The World Wide Web of Sports: A Path Model Examining How Online Gratifications and Reliance Predict Credibility of Online Sports Information

This study surveyed sports enthusiasts online to examine the impact of Internet gratifications and other factors on judgments of online credibility. Being motivated to go online for information and for entertainment were the strongest predictors of credibility of online sports sources. Reliance on traditional media and age were the only other consistent predictors of online credibility, with young, heavy media users rating the Internet as a credible source of sports information.

The World Wide Web of Sports: A Path Model Examining How Online Gratifications and Reliance Predict Credibility of Online Sports Information

As Internet research has matured, scholars have moved beyond asking who is online to address such questions as what motivates them to use the Internet and how credible do they find the information they find there. While uses and gratifications and credibility have been two main avenues of online research, little attention has been paid to fusing these two research areas even though studies suggest that motivations for media use may influence perceptions of credibility (Greenberg & Roloff, 1974).

This study surveyed sports enthusiasts online to examine the impact of Internet gratifications on judgments of "Webelievability," (the extent to which people judge the Internet as a credible source of information) (Johnson & Kaye, 2002b). This study employs a path analysis model to examine the degree to which reliance on the Web, reliance on traditional media, Web gratifications, sports knowledge and demographic variables predict credibility of sports information. More specifically, this study examines credibility of online newspapers, television sports news, radio sports news and sports magazines and the degree to which credibility of these online media for sports information is predicted by reliance on the Web, reliance on traditional media, Internet gratifications, knowledge and demographic variables.

Internet Credibility

Studies dispute whether the Internet is considered as credible as traditional news sources or even whether online news *should* be considered as credible as its traditional counterparts. For instance, while journalists increasingly rely on the Internet as an information source (Middleberg & Ross, 2001) and they judge online media sites as credible as traditionally delivered ones (Finberg & Stone, 2002), the Online News Association found that two-thirds of journalists did not feel online news met the same standards as traditional sources of news. Journalists accused online sources of being biased, unresearched and for being more prone to print rumors and gossip. Others noted that because online news sources rush to get information posted that accuracy is sacrificed. One reporter noted, "Anyone can post a Web site and call it online news... while it takes considerable time, effort and investment to establish a traditional news source" (Finberg & Stone, 2002, p. 34).

When the Online News Association questioned members of the public they were baffled why anyone would think the Internet is less credible than traditional sources. The general public believes that the print and online versions of a publication deliver the same information, just in two different forms--and online information is more up-to-date (Finberg & Stone, 2002).

Studies are similarly split on whether individuals judge the Internet as credible with several finding the Internet more (Horrigan & Rainie, 2002) or at least as credible (Johnson & Kaye, 1998, 2000, 2002a) as traditional sources, others finding that trust in the Internet is dropping (UCLA Internet Report, 2003) and still others finding neither traditional nor online sources are considered very credible (Kiousis, 2001).

Several recent studies have indicated that the Internet is emerging as a trusted source of information. For instance, Pew Research found that 69 percent of all Americans and 85 percent of Internet users expect to find reliable, up-to-date news online, and 84 percent of all Americans (and 97 percent of all Internet users) expect to find reliable information on government agency, electronic commerce, news or health care sites (Horrigan & Rainie, 2002). An earlier Pew Research study (2000) found that when Internet users are asked to compare traditional media with their online counterparts, such as CNN with CNN.com, the Internet source was judged more credible. Finally, Flanagin and Metzger (2001) discovered that Internet users went online to get information more often than they used books, magazines, televisions, newspapers, the telephone, electronic mail or interpersonal communication.

Some studies find that online sources are judged as credible as their traditional counterparts (Flanagin & Metzger, 2000; Johnson & Kaye, 1998, 2000, 2002a). For instance, Johnson & Kaye (1998, 2000, 2002a) found that while politically interested Web users distinguished between different types of media in assessing credibility (e.g. newspapers and newsmagazines were judged more credible than television news), they did not distinguish between traditional and online versions of the same media (e.g. print and online newspapers). Johnson and Kaye (2002a) argue that

Internet users are aware that media sites offer similar or, in some cases, identical content whether in online or traditional form. The difference between online and traditionally delivered media is not in what is produced, but in how that news is delivered (see also Finberg & Stone, 2002). While Johnson & Kaye (2002b) found that credibility scores had risen considerably from 1996 to 2000 because users had more experience with Internet news and were thus more able to rate its credibility, most respondents still judged online newspapers, newsmagazines and television news as only somewhat credible.

While some early studies suggested that the Internet was not considered a reliable source of information (GombdaWeb, 1998, Mashek, 1997), more recent studies have also questioned the credibility of the Internet (UCLA Internet Report, 2003). While about half of Internet users in the UCLA study consider the Internet to be reliable and accurate, the percentage judging the Internet as credible in 2002 (50.6%) fell from both 2001 (56.1%) and 2000 (52.1%). Also, just under 40 percent (39.9%) of Internet users said only about half of the information online is reliable and accurate. Nonusers, not surprisingly, put even less faith in the Internet with the plurality saying that only about half of the information online is believable. Curiously, while Internet users may not trust the information they find online, about 60 percent consider it a very or extremely important source of information with experienced users ranking the Internet higher than all other sources as an important source of information (UCLA Internet Report, 2003).

Finally, Kiousis (2001) found that none of the sources he studied—newspapers, television news or online news—rated high on credibility with three quarter of his respondents saying each source was only moderately credible. Online news may be judged as only moderately credible because Internet use is associated with low to moderate trust in government (Johnson & Kaye, in press). Internet users may not put much faith in government or major institutions, including the media.

Most of the research on Web credibility has focused on either political information (Johnson & Kaye, 1998, 2000, 2002a, 2002b) or news (Bucy, in press; Kim et al., 2001; Kiousis, 2001; Sundar, 1999). But credibility ratings may differ depending on the type of Web site examined. Flanagin and Metzger (2000) discovered that respondents judged commercial sites as significantly less credible than news, reference and entertainment ones which supports studies that have found that sources that are perceived as trying to persuade and manipulate individuals are viewed as less trustworthy (Petty & Cacioppo, 1981).

Little attention has been paid to assessing the credibility of sports information, although the scant evidence suggests sports information is considered highly credible. Gantz (1981) found that on local newscasts, sports information was rated as the most credible (86.8 percent), followed by local news (80.8 percent), state news (75.7 percent) and weather news (71.7 percent).

No studies could be found that have examined the credibility of online sports information, and studies of usage patterns of sports information offer mixed signals about whether people judge sports information as reliable. A Pew Research poll (2003) of what activities people engage in online finds that checking sports scores or information rates relatively low on the list, 23rd out of 28, although the 44 percent who said they have at one time accessed the Internet for sports information was higher than those who had gone online for political information (40 percent). On the other hand, the J.D. Powers and Associates Sports Web Site study (2001) indicated that sports sites attracted a loyal audience because they offered up-to-date information that people valued; 44 percent of user satisfaction with Websites was explained by the quality of the content rather than appearance of the site and ease of use. Similarly, the Online Publishers Association (2002) discovered that those who visited sports sites had higher-than-average brand loyalty. That is, those who access sports Web sites such as ESPN.com are more likely to get information from companion sources such as the ESPN network and magazine. The brand loyalty suggests that sports Internet users at least consider their favorite sites as credible.

Web Reliance and Internet Credibility

While past studies suggest that how often one uses a particular medium influences how credible it is seen (Austin & Dong, 1994; Wanta & Hu, 1994; Johnson & Kaye, 1998, 2000. 2002b), and that people judge their preferred medium as the most credible (Carter & Greenberg, 1965; Roper, 1977; Westley & Severin, 1964), results are mixed on whether those who more heavily rely on the Internet view it as more credible than light users.

Johnson and Kaye's study of politically interested Internet users during the 1996 election campaign found that reliance on the Web for political information was correlated with how credible they judged most online sources (Johnson & Kaye, 1998). Reliance on traditional media and the Web were the strongest predictors of credibility of online information (Johnson & Kaye, 2000).

However, Johnson and Kaye's (2002b) study of the 2000 campaign found no significant relationship between Web reliance and credibility. They speculated that because heavy users of the Internet were less likely to rely on traditional sources than light users they were not as well trained to judge which online sources were credible and which ones were not. Kiousis (2001) also failed to find a relationship between online news use and credibility.

Reliance on Traditional Media and Internet Credibility

Studies suggest that credibility is at least part a function of users' preference and familiarity with a medium. Media reliance, then, has proven one of the strongest predictors of perceptions of credibility (Greenberg, 1996 ; Rimmer & Weaver, 1987; Wanta & Hu, 1994; Westley & Severin, 1964).

Several online studies also suggest that the reliance on traditional media is the best predictor of credibility of online information, stronger than reliance on the Internet as an information source (Johnson & Kaye 1998, 2000, 2002b). For instance, Johnson & Kaye (2002b) found that reliance on traditional media was the strongest predictor of credibility of online newspapers, television news, radio news and newsmagazines. These findings suggest that the Internet supplements rather than replaces other sources of information, particularly ones that are not functionally similar to it such as newspapers, magazines and radio (Althaus & Tewksbury, 2000; Kaye & Johnson, 2003; Lin 2001a). Because heavy users of traditional media view those sources as credible, they may also see the Internet as more believable than light users (Wanta & Hu, 1994; Westley & Severin, 1964). Those who rely heavily on traditional media tend to be more media literate; they have learned how to tell biased from slanted reporting and reliable from questionable sources of information (Flanagin

& Metzger, 2000; Gilster, 1997; Potter, 1998). They bring this expertise to the Internet, using the standards they learned to judge traditional media to the online environment.

Internet Motivations and Internet Use

Uses and gratification researchers argue that audience members actively search out media messages to satisfy certain needs, and that the audience is active and goal directed (McLeod & Becker, 1981; Palmgreen, 1984). The assumptions of uses and gratifications research fits well with the Internet and its interactive features such as e-mail, instant messaging, chat and message boards that require surfers to be active users. Similarly, Internet users often actively seek out messages through search engines and links, which suggests that Web use is goal directed (Lin & Jeffres, 1998)

Different media serve different needs, and within each medium, different content can satisfy different motivations. For instance while radio music can satisfy relaxation needs (Finn 1977), television primarily serves entertainment and relaxation motivations (Perse & Courtright, 1993), while newspapers satisfy informational needs (Perse.& Courtright, 1993; Weaver & Buddenbaum, 1979). However, different content within a medium can produce different needs. For instance, while watching soap operas and game shows satisfied companionship, relaxation, passing time, arousal and habitual needs (Rubin & Rubin, 1982), those seeking to satisfy information gratifications search out television news and talk shows (Rubin & Rubin, 1982; Vincent & Basil, 1997).

Studies also suggest that different Internet content can gratify different needs. Research examining the Web in general suggest that, like television, the Web tends to satisfy entertainment, social interaction and escape needs (Charney & Greenberg, 2001; Eighmey, 1997; Ferguson & Perse, 2000; Kang & Atkin, 1999; Kaye, 1998; Lin, 2001b; Papacharissi & Rubin, 2000). However, because several functions (e.g., e-mail, chat rooms and instant messaging) and major uses of the Internet (e.g. downloading music and researching for work) do not compete functionally with other media, researchers have identified gratifications unique to the Internet such as convenience (Kaye &

Johnson, 2001; Korgaonkar & Wolin, 1999; Papacharissi & Rubin, 2000; Rosales, 2001), peer identity, communications, and coolness (Charney & Greenberg, 2001), self-disparagement, self-slighting and novel sensory outcomes (LaRose et al., 2001). On the other hand, when Korgaonkar and Wolin (1999) explored what gratifications consumers got from the business sites transaction-based security and privacy concerns and interactive control emerged as major predictors. Similarly, Kaye and Johnson (2001, 2002) found that politically interested Internet users primarily went online for guidance and information seeking while entertainment served as a weaker motivation for Web use.

Some researchers have examined how motivations for going online correlate with visiting sports sites with results differing depending on how sports sites were defined. Ferguson and Perse (2000) correlated motivations for surfing the Web with sports sites in general and found that visiting sports sites was linked to using the Web for entertainment, relaxation and social information. Lin (2001b) found that accessing sports news and information sites predicted entertainment and surveillance motivations, suggesting that those with entertainment and surveillance needs seek out sports information.

Internet Motivations and Media Reliance

Numerous studies have documented that increased motivations for using the media boosts levels of media consumption. That is, the more people are motivated to use the media to satisfy their needs, the more time they spend with the media (e.g., Greenberg, 1974; Kippax & Murray, 1980; McLeod & Becker, 1974; Rubin & Rubin, 1982). Most of these studies have been correlational, but regression analysis also indicates that motivations lead to increased use. For instance, Wenner (1983) used hierarchical regression to demonstrate that gratifications sought and obtained explained amount of exposure to network evening news and *60 Minutes*.

Research indicates that both entertainment and information motives for going online predict Internet use. Greenberg and Charney (2001) discovered that using the Internet to keep informed predicted time spent surfing the Web. In addition, four reasons for going online—to find something in particular, to pass time, to see what's going on and to use it for fun—also predicted time online. Kang and Atkin (1999) found that entertainment motives were the strongest predictors of Internet use, while Papacharissi and Rubin (2000) reported that information seeking significantly predicted Web browsing.

Internet Motivations and Credibility

While several studies have explored how amount of attention paid to the media influences credibility, little attention has been paid to how gratifications for using the media impact credibility ratings. Greenberg and Roloff (1974) argued that people are motivated to use newspapers to satisfy information needs and to use television for its entertainment content. Media consumers who are in an entertainment-processing mode are less critical and therefore less likely to search for, and discover, news errors than those in an information-processing mode and therefore would favor television news over newspapers (Greenberg & Roloff, 1974; Reagan & Zenaty, 1979). Also, Mulder (1980) discovered that the more actively people search out news the more likely they are to choose newspapers over television, suggesting that the relationship between gratifications and credibility is mediated by amount of media use. These studies suggest, then, that using the media to satisfy both entertainment and information needs may increase reliance on media which in turn will boost credibility ratings.

One study examined the relationship between motivations (convenience) for using the Internet and credibility ratings. Johnson and Kaye (2002b) explored how convenience impacted credibility ratings. Convenience proved to be the second strongest predictor of credibility perceptions of online newspapers and online radio news and was the third strongest indicator of online television news and news magazines.

Demographics, Knowledge and Internet Credibility

Studies of traditional media suggest, paradoxically, that those individuals who most heavily rely on the media—highly educated males with high incomes—are the most critical of media content (Mulder, 1981; Robinson & Kohut, 1998).

Research on online media also suggest that women, the young, and those with lower incomes and education deem the Internet as believable (Bucy, in press; Finberg & Stone, 2002;

Johnson & Kaye, 1998, 2000a), although differences sometimes disappear after controlling for other factors (Johnson & Kaye, 2000, 2002b).

Johnson and Kaye (2002b) speculated that because more women, senior citizens and those of high socioeconomic status are going online (Howard et al., 2001; Nie & Ebring, 2000; UCLA Internet Report, 2003) that demographics would have less of an impact on credibility judgments. However, they found in a study exploring the effects of Internet experience on credibility (Johnson & Kaye, in press) that demographics were some of the most significant predictors of online credibility. As previous studies suggested, younger individuals with less education judged the Internet as more credible. However, men rated all online media as more credible than women, which went against previous studies of traditional and online media. Johnson and Kaye (in press) suggested because research has found that men have been online longer than women, spend more time online and engage in more activities (Howard et al., 2001) that experience online may have helped men judge better which sites to believe and which ones to discount.

Gender could also influence credibility ratings for sports sites. Studies suggest that men visit sports sites considerably more than females (Cyber Dialogue, 2001; eMarketer, 2002; Screen Digest, 2000) as well as discuss sports more often in online groups (Pew Research, 2001); indeed, gender has emerged as one of the strongest predictors of visiting sports sites (Ferguson & Perse, 2000). Studies suggest that those who visit sports sites are younger than average (e.Marketer, 2002; Screen Digest, 2000), are affluent (Beard, 2002; Cyber Dialogue, 1999; Screen Digest, 2000) yet have less education (Cyber Dialogue, 2001) than the average user. Among young males, sports sites are the second most popular use behind using search engines (eMarketer, 2002).

Prior knowledge can also influence credibility judgments. Those who are highly knowledgeable about a topic have more ability and motivation to process information and are more able to evaluate it critically (Eagle & Chaiken, 1993; Eastin, 2001; Wathen & Berkell, 2002). Indeed, Eastin (2001) found that those more knowledgeable about health-related information judged health sites more credible than the less knowledgeable.

Research Questions and Hypotheses

This study examines the credibility of online media for sports information and poses the following research question and hypotheses.

RQ1. To what degree will individuals who regularly use the Internet for sports news judge online newspaper sports, online broadcast and cable television sports news, online radio sports news, and online sports magazines as credible?

Based on the path analysis model, the following relationships are hypothesized:

H1: The credibility of online newspapers, television news, radio news and newsmagazines will be positively predicted by a) reliance on the Web, b) reliance on traditional media, c) using the Web to satisfy informational and entertainment gratifications, d) sports knowledge, and e) demographics (gender, age, income, and education).

H2: Reliance on the Web will be positively predicted by a) reliance on traditional media, b) Web gratifications, c) sports knowledge, and d) demographics.

H3: Reliance on traditional sources of sports news will be positively predicted by a) Web gratifications, b) sports knowledge, and c) demographics.

H4: Web use gratifications will be positively predicted by a) sports knowledge, and b) demographics.

Method

A survey measuring perceptions of credibility of online and traditional media sources of sports information was posted on the World Wide Web from December 13, 2002 - March 5, 2003. The survey was systematically promoted to Internet users who access sports information online. Announcements informing online users about the survey were sent to media and sports- oriented Websites,[1] newsgroups,[2] electronic mailing lists,[3] and bulletin boards.[4] Additionally, the survey site URL was submitted directly to search engines such as Google and ExactSeek as well as to sites that automatically post the URL on numerous search engines.[5] Finally, we used a "snowball" technique where the authors sent announcements to sports fans in their e-mail address book and asked them to fill out the survey and to send it to other sports fans (Babbie, 2001). Additionally, the survey was set up in such a way that respondents who wished to could automatically forward the survey to other sports fans to fill out.

The survey specifically targeted Internet users who go online for sports information. Generating a random sample of this group is next to impossible because the Internet does not have a central registry of users and so unlike telephone and mail surveys, samples cannot be produced through census lists or random digit dialing type techniques such as random e-mail generators (Kaye & Johnson, 1999; Witte, Amoroso, & Howard, 2000; Wu & Weaver, 1997). Therefore, this study employs a convenience sample of sports- interested Internet users.

Dependent Measures

Past studies suggest judgments of media credibility are influenced by how credibility is defined (Gaziano & McGrath, 1986). Although media credibility has been measured in several ways, believability, fairness, accuracy, and depth or completeness are four common attributes that have consistently emerged from past studies that created credibility indexes (Gaziano & McGrath, 1986; Johnson & Kaye, 1998, 2000, 2002a).

Independent Measures

Source reliance. Respondents were asked to assess their levels of reliance on traditional media (broadcast and cable television news, radio news, newspapers, and sport magazines). The five-point scale ranged from "heavily rely on" to "don't rely at all."

Web Gratifications. As part of the path model proposed by this study, motivations for going online for sports information are hypothesized to predict reliance on traditional media, Web reliance and perceptions of online credibility. Factor analysis was conducted on 20 items assessing respondents' reasons for turning to online sources for sports information. Possible responses ranged from (1) strongly disagree to (5) strongly agree. The items were adapted from other studies that examined uses and gratifications of accessing the Internet (Kaye, 1998; Kaye & Johnson, 2001, 2002).

Sports Literacy

Respondents were asked to rate themselves on their knowledge of sports. The options were "very literate," "literate," "somewhat literate," "barely literate," and "not at all literate."

Demographics

This study examined associations between perceptions of credibility of online sources and gender, age, income, and education.

Data Analysis.

First, credibility indexes for each online medium (newspaper sport news, broadcast and cable television sport news, radio sport news sites, and sport magazine sites) were created from a series of questions examining judgments of accuracy, fairness, believability, and depth of each online source. Then frequencies were run on the five online credibility indexes.

Next, the 20 motivation items were then factored by principal components analysis with varimax rotation. Items whose primary loadings were greater than .45 were assigned to a particular factor (Stevens, 1986).

Lastly, a series of hierarchical multiple regressions were conducted to test whether perceptions of online credibility can be explained by reliance on traditional and online media, Web use motivations, sports literacy, and demographics. Regressions tested the path model for each online medium - newspapers sport news, broadcast and cable television sports news sites, radio sport news sites, and sport magazines. For example, to test the model for online sports magazines, the dependent variable was credibility of that medium. The independent variables entered into the equation were reliance on the Web, reliance on traditional sports magazines, Web use motivations, sports knowledge and demographics (gender, age, education, income).

Results

Respondent Profile

The online sport survey was completed by 840 respondents who connect to online sports information. The respondents spend 6.1 hours per week seeking online sports information and 11.7 hours on the Web in general. Almost three-quarters (74.6%) consider themselves literate or very literate in matters of sports and they devote an average of 15.8 hours each week to playing, watching, reading or learning about sports.

More males (56.7%) than females (43.3%) completed the survey. Respondents average 29 years and 4 months of age, and almost all (86.3%) are white, 4.3% are Hispanic, 3.5% are Black and the remaining classified themselves as other non-whites. The respondents are highly educated with just more than nine out of ten (93.4%) reporting some college, a college degree or higher. Just over 6 out of 10 (61.2%) report an annual income of less than $25,000, slightly more than one-quarter (26.0%) fall into a middle income range between $25,000-65,000, and the remaining 12.7% earn more than $65,000 per year. While this is a convenience sample, demographic comparisons with other Internet surveys suggest the sample may be fairly representative of the Internet population at the time of data collection in 2002, at least in terms of gender and age (Cyber Dialog, 2001; eMarketer, 2002; Screen Digest, 2000)

Credibility of Online Sources

Anyone can post whatever information they want to online and this lack of accountability calls the Web's credibility into question. The first research question assesses respondents' judgments of credibility of online newspaper sports, broadcast and cable television sports pages, radio sports sites, and sports magazines sites. Overall, most people judged online information as moderately credible, and between 19.3% and 38.7% rated all the online media, except online radio, as very credible. Online cable news sports has the largest percentage (75.6%) of respondents judging it as "moderately" or "very" credible, closely followed by online newspaper sports sites (70.8%) (Table 1).

Online Television News. Cable television news sites, such as ESPN.com and CNN.com, has the largest percentage of respondents rating it as very credible (38.7%). On the other hand, broadcast television (i.e. ABC.com) trailed all sources but online radio in credibility with 19.3% rating it very credible.

Newspapers and Sports Magazines Online: Both online newspaper sports and sports magazines sites are considered highly credible sources. About two-thirds of respondents deem both sources as "moderately" or "very" credible (70.8% newspapers, 65.6% sports magazines).

Radio News Online: Radio news has the smallest percentage of respondents (48.9%) rating it as moderately to very credible. It also has the highest percentage (11.5%) of those who feel it is "not at all" or "not very" credible. Audio sports news rates lower in credibility than any other online medium (Table 1).

Motivations for Using Online Sources for Sports Information

Principal components analysis was conducted on 20 the motivation items yielding three distinct factors: Information Seeking (eigenvalue, 7.6; variance explained, 38.3%), Entertainment (eigenvalue, 2.2; variance explained, 11.0%), Gambling (eigenvalue, 1.2; variance explained, 6.1%). Reliabilities of the factors for each factor range from .73 to .88 (Table 2).

The eight measures that loaded on the information-seeking factor indicate that individuals use the Internet to quickly get up-to-the-minute information that they cannot get anywhere else. The entertainment factor includes eight variables that indicate the Internet is used as a source of fun, relaxation and as a topic of discussion with others. Finally, gambling includes three variables that reflect a utilitarian use of the Internet for fantasy sports leagues, the office pool and to use in arguments with others (Table 2).

Predictors of Online Media Credibility

This study specifically poses a model that predicts that demographic and sports literacy will lead to motivations for using the Web for sports information, which in turn will lead to reliance on traditional media, which will lead to reliance on the Web and to online source credibility (Figure 1).

Hypothesis One. This hypothesis predicts that credibility of the Internet would be explained by a) reliance on the Web, b) reliance on traditional media, c) Web motivations d) sports knowledge and e) demographics. Indeed, traditional media use, two of the three Web gratifications, and certain demographic variables predict online credibility.

Only Web reliance was completely unrelated to perceptions of credibility of online sports information - none of the relationships were significant (Table 3). Therefore, Hypothesis 1a is not supported.

Credibility of all the online media under study is weakly predicted by reliance on the traditional form of that medium: online cable television sports (beta=.12, $p<.01$), online radio sports news (beta=.12, $p<.01$), online sports magazines (beta=.11, $p<.01$), online newspaper sports (beta=.10, $p<.01$), and online broadcast television sports (beta=.10, $p<.01$) (Table 3). Hypothesis 1b is consistently though weakly supported—the more respondents rely on a particular traditionally delivered medium the more they judge that medium as credible.

Hypothesis 1c is partially supported as using the Web for information seeking reasons predicts credibility of all but online radio news, while using the Web for entertainment predicts the credibility of online radio sports, online broadcast television sports and online newspaper sports. Using the Web for gambling does not predict credibility of online media. Using the Web for seeking online sports information, such as sports scores and team standings, most strongly predicts the credibility of online cable television sports (beta=.26, $p<.001$), closely followed by online sports magazines (beta=.21, $p<.001$), and online newspaper sports (beta=.21, $p<.001$). Using the Web for entertainment purposes, such as judging who will win or for excitement, predicts the credibility of online radio sports news (beta=.18, $p<.01$), broadcast television sports (beta=.15, $p<.01$) and online newspapers (beta=.11, $p<.05$). The findings suggest that both online newspaper sports and online broadcast television sports are accessed for both information seeking and entertainment reasons, but when respondents seek more serious sports information they turn to online cable television sports and online sports magazines and leave being entertained to online radio sports (Table 3).

Hypothesis 1d is not supported as the degree of sports knowledge did not significantly predict four of the five variables, and the only significant relationship is negative. Lower degrees of sports literacy lead to higher judgments of the credibility of online newspaper sports (beta=-.09, p<.05). Those respondents who claim they are not very sports knowledgeable deem online newspaper sports more credible than the other online sports sources (Table 3).

In some cases demographics are significant predictors of credibility of online sports media, so hypothesis 1e is partially supported. Age negatively predicts the credibility of all online media. Younger respondents are more likely to assess online sports information as more credible than older respondents. Education and income only predict the credibility of online newspaper sports. Those with higher levels of education (beta=.15, p<.01) but lower incomes (beta=.-11, p<.05) view online newspapers as credible. Gender does not predict the credibility of any of the online sports information sources.

Hypothesis Two - Predictors of Reliance on the Web

This study hypothesized that reliance on the Web for sports information would be predicted by a) reliance on traditional media, b) motivations for going online, c) sports knowledge and d) demographic variables. All components of this hypothesis were supported or at least partly supported. Reliance on two of the five traditional media predicted Web reliance as did two of the three Internet motives. Sports knowledge influenced Web reliance as did two of the four demographic variables.

Reliance on the Web is negatively predicted by reliance on broadcast television sports (beta=-.12, p<.01) and positively predicted by traditional sports magazines (beta=.07, p<.05), therefore Hypothesis 2a is partially supported. Those who rely on over-the-air television are less likely to rely on the Web for sports information, however those who rely on sports magazines are more likely to rely on the Web, suggesting the Web may supplement magazines (Table 4).

Reliance on the Web is also predicted by motives for going online. Therefore, Hypothesis 2b is supported. Information seeking is a strong, positive (beta=.43, p<.001) predictor of Web reliance. Entertainment is a weaker, positive predictor (beta=.11, p<.01) but going online to gamble is not related to reliance. Sports enthusiasts who are motivated to go online for information and for entertainment tend to develop a reliance on the medium (Table 4).

Knowledge about the world of sports positively predicts (beta=.18, p<.01) Web reliance, which supports Hypothesis 2c. Age and education are also positive predictors of reliance on online information (beta=.08, p<.05; beta=.07, p<.05, respectively), which partially supports Hypothesis 2d. Older, better educated respondents who report great depth of sports literacy rely more on the Web for sports information than do those who are younger, not as educated or sports literate. However, gender and income are unrelated to reliance on the Web for sports information (Table 4).

Hypothesis Three - Predictors of Reliance on Traditional Media

Hypothesis 3 predicts that reliance on traditional sources of sports information will be positively predicted by a) motivations for using the Web, b) sports literacy and c) demographics. Sports knowledge and entertainment gratifications were the strongest predictors of relying on traditional sources for sports information, although gender and age influenced reliance on certain traditional news sources. Being motivated to going online for information and gambling as well as socioeconomic status were unrelated to relying upon traditional sources of sports information

Hypothesis 3a is partially supported as going online for entertainment positively predicts reliance on traditional sports magazines (beta=.23, p<.001), over-the-air radio sports (beta=.15, p<.01), broadcast television news (beta=.14, p<.01) and cable television news (beta=.13, p<.01). Neither going online for sports information nor gambling is related to reliance on traditional news sources. The findings suggest that those who go online seeking entertainment may also judge television, radio, and sports magazines, but not newspaper sports sections, as delivering entertaining sports fare.

Sports knowledge, gender, and age are also related to reliance on traditional media for sports information. Those who report greater knowledge about sports in general tend to rely on all of the traditional sources of sports information. Younger respondents rely on broadcast television news, cable television news and sports newsmagazines more than older individuals. Results for gender are split. Males are more likely to rely on broadcast news for sports information while women sports fans choose radio news. Education and income did not predict traditional media reliance.

Therefore, while Hypothesis 3b was supported, Hypothesis 3c was only partially supported (Table 5).

Hypothesis 4: Predictors of Motives for Going Online for Sports Information

The final hypothesis claims that sports knowledge and demographics positively predict reasons for going online. Hypothesis 4a is supported as each Web use motivation is positively predicted by sports literacy: information seeking (beta=.33, p<.001); entertainment (beta=.20, p<.001) and gambling (beta=.12, p<.01).

Hypothesis 4b is partially supported as several demographic characteristics predict reasons for going online. Entertainment is negatively predicted by education (beta=-.17, p<.001) suggesting that those who are knowledgeable about sports but have lower levels of education are more likely to go online for sports entertainment. Those who go online for sports gambling are more likely to be younger (beta=-.22, p<.001), sports literate(beta=-.12, p<.001) females (beta=-.17, p<.001) with low levels of education (beta=-.15, p<.001) (Table 6).

Discussion

This study surveyed sports enthusiasts online to examine the extent to which they viewed Internet information as credible as well as the influence of Internet gratifications on the believability of online information. More specifically, this study employed a path analysis model to examine the degree to which reliance on the Web, reliance on the traditional media, Web gratifications, sports knowledge and demographic variables predict the credibility of sports information.

Credibility of Online Media

Past studies, which have looked largely at the credibility of news or political information, have disagreed on the extent to which the Internet is viewed as a reliable source of information, and little attention has been paid to how credible sports fans judge online information. This study suggests that online sports information is considered highly credible, which supports an earlier study of traditional media (Gantz, 1991). While the plurality judged the online sports information as moderately credible, the percentage of those who rated the different online sources as moderately or very credible outdistanced those who judged news and political information as credible. For instance, More than 7 in 10 respondents judged online newspapers moderately or very credible for sports information and 65.6 percent said the same of online sports magazines. In comparison, only 40 percent in the Johnson and Kaye study (2002a) of political information thought online newspapers were a credible source and 37.3 percent rated online newsmagazines as moderately or very credible. Kiousis (2001) found that three-quarters of his respondents judged traditional and online sources as only moderately credible.

Studies indicate that the most important factors that influence credibility of an online message are accuracy, comprehensiveness, fairness and currency (Alexander & Tate, 1999; Finberg & Stone, 2002). The J.D. Power and Associates study (2001) suggest that Internet sports sites rate highly on these criteria, and sports fans rely on the Web for basic information such as scores and schedules that is likely to be highly accurate and therefore highly credible.

This study supports earlier works that find that the public do not judge each online medium as equally credible, but distinguish between various sources. Print sources (online newspapers and sports magazines) tended to score higher than broadcast sources (online radio and broadcast television sites) which reinforces other studies that suggest people look to print sources for more in-depth news (Chew, 1994; Culbertson & Stempel III, 1986). However, cable television sites like ESPN.com emerged as the most credible, which is in line with previous studies in which cable television websites were ranked as the most believable (Finberg & Stone, 2002; Pew Research, 2000). The credibility scores paralleled scores for reliance. Sports fans said that other than the Web (74 percent), they relied on cable television (67 percent) the most for their sports news and information followed by newspapers (53 percent). This reinforces earlier studies that indicate that credibility of a medium is strongly related to how often one uses it (Austin & Dong, 1995; Wanta & Hu, 1994; Johnson & Kaye, 1998, 2000, 2002a) and that people judge their preferred medium as the most credible (Roper, 1977; Westley & Severin, 1964).

This study suggests, then, that those studies (e.g. The UCLA Report, 2003) that simply ask people to rate the credibility of the Internet are off base. Credibility ratings differ depending on the medium and the message and, as other studies suggest, the types of Websites studied (Alexander & Tate, 1999; Flanagin & Metzger, 2000).

Credibility and Web Reliance

Web reliance did not significantly predict online credibility, which supports recent studies examining credibility of political information (Johnson & Kaye, 2002b; Kiousis, 2001). Reliance on traditional sources did not predict reliance on the Web. Those who read traditional sports magazines are weakly, though significantly likely to go online while reliance on broadcast news was negatively linked to using the Web. These sports fans may be motivated to go online for both news and entertainment and they may be quite knowledgeable about sports, but because of their limited experience with traditional media they may be less trained to know what online sources are credible and which ones are not (Flanagin & Metzger, 2000; Gilster, 1997; Potter, 1998).

Credibility and Reliance on Traditional Media

Reliance on traditional media has proved to be the strongest predictor of online credibility in previous studies. However, media reliance served as a weak, though consistent predictor of online credibility, significantly predicting credibility of all five online sources. However, it trailed online motives and age as credibility predictors. Johnson and Kaye (2002b) argue that traditional media is a strong predictor of media credibility because those who rely on traditional sources are more media literate. They have considerable experience with the media and are thus able to better tell objective from slanted reporting and what sources of information are reliable. However, their average age in their survey was 42 compared to 29 for this study. Online sports enthusiasts, because of their age, may have less experience with traditional media than politically interested users and are therefore less able to judge its credibility.

Credibility and Motivations

Credibility researchers argue that different individuals have different needs. Consequently, studies have found that those who use the Internet for general information needs (Charney & Greenberg, 2001; Eighmey, 1997; Ferguson & Perse, 2000; Kang & Atkin, 1999; Kaye, 1998; Lin, 2001b; Papacharissi & Rubin, 2000) differ from those who rely on it for political (Kaye &

Johnson, 2001, 2002) or commercial reasons (Korgaonkar & Wolin, 1999). Studies have also suggested that because of the unique, interactive nature of the Internet that it serves some unique needs such as convenience. This study found that like politically interested users, sports fans primarily use the Internet for information. However, like studies of general Internet use, entertainment also served as a major motive of going online, demonstrating that different audiences indeed have different needs. While politically interested Internet users are primarily interested in discovering information, sports enthusiasts have more varied needs. They are motivated to going online for information such as game scores, schedules and sports headlines, but because sports is entertainment they are also motivated to go online to watch or listen to a game or discuss the fortunes of their favorite team with other fans or for other reasons of entertainment. This study also found a unique motivation for going online: sports gambling. However, gambling was the weakest of the three factors and was not related to credibility or reliance on either traditional or online media.

Credibility and uses and gratifications have emerged as two of the most popular areas of research for Internet communication, but little attention has been paid to exploring the degree to which motivations for using the media predict credibility ratings. This study found that two Internet motives, information and entertainment, proved to be the strongest predictors of online credibility. In particular, being motivated to go online to seek information was the strongest predictor of credibility of online newspapers, cable news, broadcast news and news magazines while entertainment was the second strongest predictor of online newspapers, online broadcast television news and online radio news credibility. On the other hand, the third gratification, gambling, was unrelated to credibility ratings. While several studies have documented that reliance on the media influences credibility ratings, this study asserts that researchers also need to explore motivations for going online. Those who are motivated to either to seek out information or to be entertained by sports information are more likely to judge it credible.

Greenberg and Roloff (1974) and Mulder (1980) had suggested that motivations for using the media directly influenced credibility ratings and indirectly through their relationship with reliance. Those who were motivated to use the media would rely on it more and thus judge it as more credible. However, this study found that while being motivated to search for information was the strongest predictor of Web reliance and that entertainment weakly predicted Web reliance, Web reliance did not predict credibility. And while entertainment motives predicted four of the five traditional news sources which also predicted credibility, being motivated to search for information did not predict reliance on traditional news sources. Therefore while this study found a direct influence of motivations on credibility, the relationship between motivations and credibility do not appear to be mediated by online or traditional media reliance.

Researchers have debated whether the Internet is serving as a substitute for traditional media (Cai, 2002; Kaye, 1998; Lin 2001a) or whether it is supplementing media use (Kaye & Johnson, 2003). This study suggests that for sports enthusiasts, that traditional media is complementing rather than replacing traditional media. J.D. Powers and associates found that sports fans are attracted to the Internet because it is a quick and convenient source of sports information such as game results, sports headlines, standings and schedules and that online information is up to date (J.D. Powers, 2001). However, while it is possible to watch a webcast or listen to online radio of a sports event, those who are motivated to use the Web for entertainment still turn to the traditional media to view games, watch sports shows like *Sports Center* and listen to sports talk radio to relax and for entertainment and excitement.

Credibility and Demographics

Previous studies suggest that the young, women and those of lower socioeconomic status—a group that traditionally reports low traditional media readership—are the most likely to judge the Internet as a credible source of information (Bucy, in press; Finberg & Stone, 2002; Johnson & Kaye, 1998, 2000a, in press), although the influence of demographics sometimes disappears after controlling for other factors (Johnson & Kaye. 2000a, 2002b). This study found that other than age, demographics had little influence on credibility ratings. Young sports fans were more likely to judge all information sources as credible and college students (those young, with high education and low income) also rate online newspapers as credible. No other demographic variables were significant. Surprisingly, gender had little influence on credibility although earlier studies suggested that males were predominantly more likely than females to visit sports sites (CyberDialogue 2001; eMarketer, 2002; Ferguson & Perse, 2000; Screen Digest, 2000). However, while the majority of users in this survey were male, they hardly overwhelmed the female population; 43 percent of users were female. Results, then, support the claims of Johnson & Kaye, 2002b) that because the Internet is becoming more mainstream as more women, senior citizens and those of low socioeconomic status are going online (Howard et al., 2001; Nie Ebring, 2000; UCLA Internet Report, 2003), that demographics will have less influence on credibility ratings.

Credibility and Sports Knowledge

Sports knowledge was the major predictor of both information and entertainment motives for going online and was the second strongest predictor of Web reliance. Sports knowledge was also the most consistent predictor of reliance on traditional sources, which suggests that those who are knowledgeable about sports have a greater interest in sports and thus are more likely to be motivated to using the Internet and other sources to find sports information. However, sports knowledge was negatively linked to credibility of online newspapers and unrelated to other sources of information, which runs counter to the study by Eastin (2001) on health information that more knowledgeable individuals were more likely to judge Internet information as credible. However, studies have found that those who are highly knowledgeable about a topic have more ability and motivation to process information and they are more able to evaluate that information critically (Eagle & Chaiken, 1993; Eastin, 2001; Wathen & Berkell, 2002). Therefore, the sports knowledgeable may be more aware of errors and biased reporting in online newspapers than less knowledgeable individuals and therefore less likely to judge that information as credible.

Limitations

This study relies on a convenience sample of Internet sports users so that results cannot be generalized to the entire population or even to the population of online sports users. Reaching this particular subset of users is challenging because the Internet does not conform to traditional methods of data collection. However, scholars recognize that in situations where probability sampling is not possible, that non-probability sampling is acceptable (Babbie, 2001). This study posted announcements of the survey on sites where sports fans are likely to visit: sports listservs and usenet groups, sports bulletin boards and it was linked to the sports pages of media sites. We also used a snowball technique where the survey was sent to sports fans and they were asked to send it to other sports fans. Our survey included an option for respondents to send the survey to other sports fans. Researchers suggest that careful use of this sampling technique can produce samples that be representative of a specific subset of the population, like sports fans (Babbie, 2001). Comparisons with other studies of online sports users suggest that this population may be representative in terms of age and gender (Cyber Dialogue, 2001; eMarketer, 2002; Screen Digest, 2000), but may be more educated and less affluent than earlier research (Beard, 2002; Cyber Dialogue, 1999; Screen Digest, 2000).

This study of online sports fans found that enthusiasts rate the Internet as more credible than previous studies of general news and political information. This was one of the first studies to examine the connection between credibility and uses and gratifications and found that motives for using the Internet indeed predicted credibility of online information.

Little attention has been paid to how sports enthusiasts use the Internet. Future studies could explore other questions such as what activities do sports enthusiasts engage in online and is the Internet serving as a substitute for more traditional sources. This study found that sports fans judged the Internet as more credible than politically interested and general Internet users. This study was limited to credibility of sports information. Future studies could examine whether sports fans judge sports entertainment and sports merchandising/apparel sites as any less credible than information sites. Further research could also explore other groups who go online mainly for entertainment (e.g. gaming and downloading music) or for commercial information to see if they view the Internet as any more or less credible than sports fans.

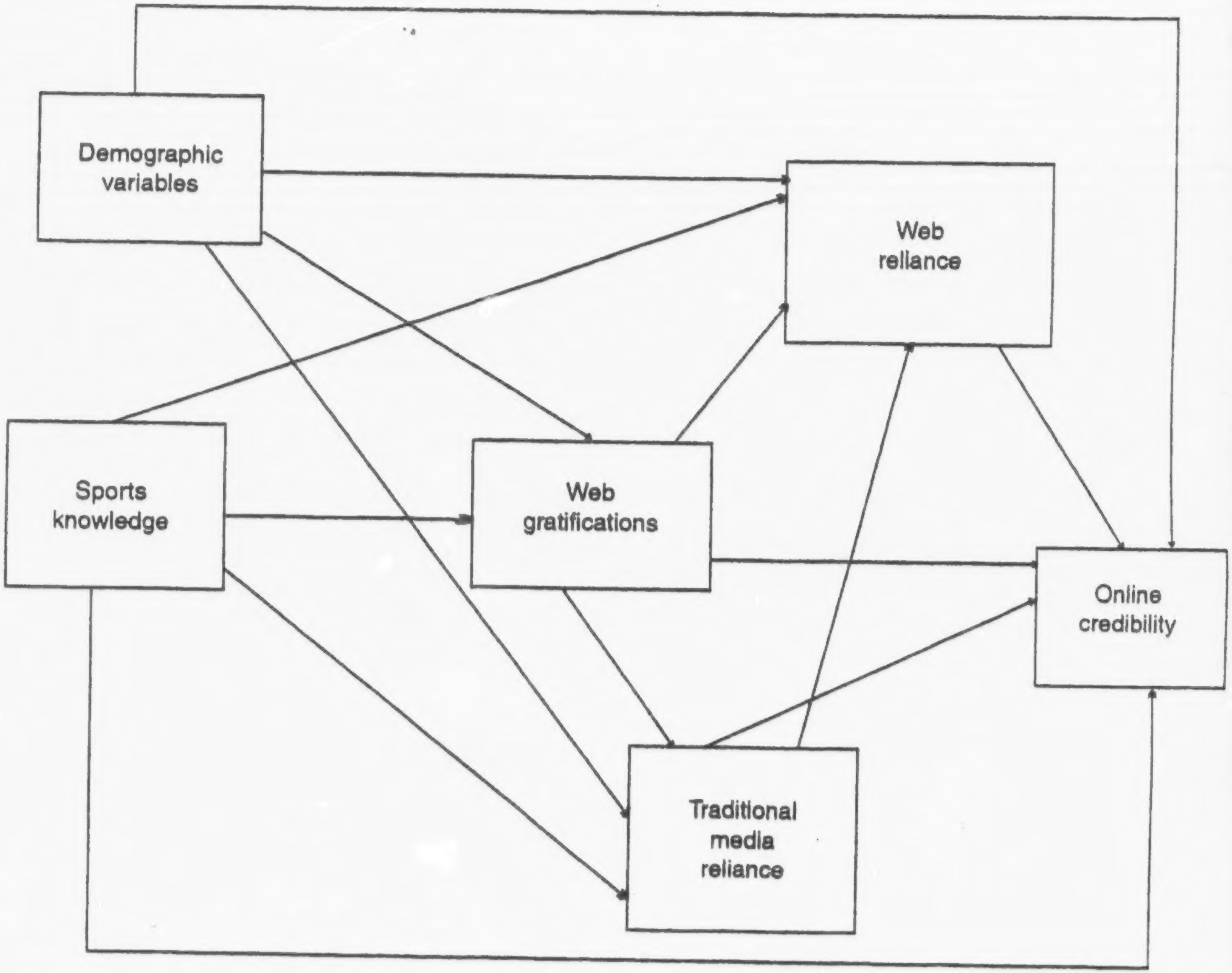

Demographic variables
Web reliance
Sports knowledge
Web gratifications
Online credibility
Traditional media reliance

Figure 1: Path model of predictors of online credility of sports news

TABLE 1

Perceptions of Credibility and Depth of Online Sources for Sports Information

2003 Credibility
(Mean Scores as percentages)

Online Newspapers (N=810)	
Not at all…	.7
Not very…	3.1
Somewhat...	25.4
Moderately…	43.8
Very....	27.0

Online Magazines (N=794)	
Not at all…	1.1
Not very…	3.5
Somewhat...	29.7
Moderately…	39.0
Very....	26.6

Online Broadcast Television (N=798)	
Not at all…	1.1
Not very…	6.5
Somewhat...	32.5
Moderately…	40.6
Very....	19.3

Online Cable Television (N=802)	
Not at all…	.8
Not very…	2.4
Somewhat...	21.2
Moderately…	36.9
Very....	38.7

Online Radio (N=782)	
Not at all…	3.5
Not very…	8.0
Somewhat...	39.6
Moderately…	33.8
Very....	15.1

Table 2

Motivations for Using the Web for Sports Information
Factor Matrix

World Wide Web Use Motivations	F1 Information	F2 Entertainment	F3 Gambling
"I use the Web for sports information..."			
Factor 1 Information Seeking			
to access sports information quickly	.83		
to find specific sports information	.82		
because sports information is easy to obtain	.77		
for up-to-the minute general sports information	.75		
for sports information I can't get elsewhere	.70		
to see up-to-the minute sports scores	.67		
to keep up with main sports issues of the day	.55		
to see team standings	.49		
Factor 2 Entertainment			
because the Web is entertaining		.74	
because the Web is exciting		.73	
to remind me of my team's strong points		.70	
to enjoy the excitement of a game		.66	
to judge who will win		.62	
to talk sports with others		.58	
to judge qualities of teams and players		.57	
because the Web helps me relax		.58	
Factor 3 Gambling			
to do well in my fantasy sports league			.84
to do well in a bet or office pool			.84
to use as ammunition in arguments with others			.54
Eigenvalue	7.6	2.2	1.2
Variance Explained	38.3	11.0	6.1

The item "for unbiased viewpoints' did not load on any factor and thus is excluded from analysis.

TABLE 3
Predictors (Beta Weights) of Credibility of Online Sports Information

Variable	Online Newspapers	Online Broadcast TV News	Online Cable TV News	Online Radio News	Online Sports Newsmagazines
Web reliance	-.01	-.04	-.00	-.08	.04
Traditional media reliance	.10**	.10**	.12**	.12**	.11**
Information motive	.21***	.14**	.26***	.08	.21***
Entertainment motive	.11*	.15**	.04	.18**	.03
Gambling motive	.02	.00	.02	-.06	-.01
Sports knowledge	-.09*	-.02	.01	-.06	-.04
Gender	.00	-.03	-.04	-.04	.00
Age	-.11*	-.16**	-.24***	-.13*	-.14**
Education	.15**	.01	.07	.01	.08
Income	-.11*	-.06	-.05	-.08	-.08

*p.>05, **p.<.01, ***p.<001

TABLE 4
Predictors (Beta Weights) of Web Reliance

Variable	Web Reliance
Reliance On:	
Traditional Newspaper Sports News	-.00
Traditional Broadcast TV Sports News	-.12***
Traditional Cable TV Sports News	.04
Traditional Radio Sports News	.05
Traditional Sports Magazines	.07*
Information Motive	.43***
Entertainment Motive	.11**
Gambling Motive	-.05
Sports Knowledge	.18***
Gender	-.01
Age	.08*
Education	.07*
Income	.02

*p.>05, **p.<.01, ***p.<001

TABLE 5
Predictors (Beta Weights) of Reliance on Traditional News Sources

Variable	Newspapers	Broadcast TV News	Cable TV News	Radio News	Sports Newsmagazines
Information motive	-.02	-.07	.07	-.05	-.02
Entertainment motive	.06	.14**	.13**	.15**	.23***
Gambling motive	.00	-.00	.06	.06	.03
Sports knowledge	.21***	.09*	.26***	.14**	.18***
Gender	.02	.12**	-.03	-.09*	.02
Age	-.04	-.10*	-.21***	.02	-.21***
Education	.01	-.01	.06	-.00	-.04
Income	.00	-.03	-.04	.01	.02

*p.>05, **p.<.01, ***p.<001

TABLE 6
Predictors (Beta Weights) of Motives for Using Online Sports Information

Variable	Information Seeking	Entertainment	Gambling
Sports knowledge	.33***	.20***	.12**
Gender	-.06	.00	-.17***
Age	.00	-.09	-.22***
Education	-.04	-.17***	-.15***
Income	.01	.03	.01

*p.>05, **p.<.01, ***p.<001

Notes

1. We were provided links to the sports sections of the following media and sports sites: Daily Egyptian (Southern Illinois University), Daily Collegian (Penn St.), Daily Beacon (University of Tennessee), DoAne (Neb.) Owl, Red and Black (University of Georgia), georgiadogs (University of Georgia athletics website), WPSD-TV (Paducah, KY) and WBIR-TV (Knoxville).

2. We posted to Usenet groups devoted to professional and college sports (both men and women) and to specific team sites. We tried to cover a wide range of sports: football, basketball, baseball, soccer, hockey, tennis, golf, auto racing, pro wrestling, figure skating, gymnastics, extreme sports, and track. We also posted to fantasy league, memorabilia and collecting, sports video game and gambling Usenet groups:

alt.sports.football.arena, alt.sports.football.pro, rec.sport.football.pro, rec.sport.football.fantasy, rec.collecting.sport.football, red.sport.football.college, alt.sports.football.pro.gb-packers. alt.sports.football.pro.dallas-cowboys, alt.sports.football.pro.buffalo-bills, alt.sports.football.pro.chicago-bears, alt.sports.football.pro.sf-49ers, alt.sports.football.pro.oak-raiders, rec.sport.basketball.pro, rec.gambling.sports, rec.collecting.sport.basketball, alt.sport.basketball.pro.fantasy, comp.sys.ibm.pc.games.sports, alt.sports.basketball.nba, rec.sport.basketball.women, rec.sport.basketball.college, alt.sports.basketball.nba.la-lakers, alt.sports.basketball.nba.sac-kings, alt.sports.basketball.nba.dallas-mavs, alt.sports.basketball.nba.boston-celtics, alt.golf.forsale, rec.sport.golf, rec.sport.baseball, rec.sport.baseball.fantasy, rec.sport.baseball.college, alt.sports.baseball, alt.sports.baseball.oakland-A's, alt.sports.baseball.atlanta-braves, alt.sports.baseball.bos-redsox, alt.sports.baseball. ny-yankees, alt.sports.baseball.ny-mets, alt.sports.baseball.stl-cardinals, alt.sports.baseball.chicago-cubs, alt.sports.baseball.la-dodgers, rec.sport.soccer, alt.sports.soccer, rec.autos.sport.fl, rec.autos.sport.info, rec.autos.marketplace, rec.collecting.sport.misc, rec.auto.sport.nascar, rec.autos.sport.indy, rec.autos.sport.cart, alt.sports.hockey.nhl, alt.sports.hockey.fantasy, alt.sports.hockey.nhl. det-redwings,alt.sports.hockey.nhl.buffalo-sabres, alt.sports.hockey.nhl.dallas-stars, alt.sports.hockey.nhl.sj-sharks, alt.sports.hockey.nhl.phila-flyers, alt.sports.hockey.nhl.pit-penguins, alt.sports.hockey.nhl.tor-mapleleafs, rec.sport.hockey.field, rec.sport.tennis, alt.tennis, rec.sport.pro-wrestling, rec.sport.skating.ice.figure, alt.skake.figure, alt.sports.gymnastics, rec.skydiving, rec.climbing, rec.sport.paintball, alt.sport.track-field, rec.running

3.Messages were sent to the following sports and sports team listservs, as well as a few listservs that the authors were subscribed to:

sports management listserv, MCMAGRAD-L, arenafootball, tampabaystormarenafootball, kennelclub, collegebcs, nfl2day2, vikings, nflonline, worldofthenfl, chicagobeargrinandbearit, buffalobillsfanclub, dallascowboyfanclub, packerplace2, miamidolphinsfootball, go49ers, raidernation2, anyjetslounge, clevelandbrownsclub, nba2day, thenbahouse, wnbaplayersclub, newlakersbestfans, sacramentokingscourt, maverickfansportsbar, bostonculticfans, knicks2, baseballfansunite, mlbuniverse, mlbtoday, baseballdatabank, AMAZING_Jeter, oaklandasdynamictrio, oaklandas, atlanta_braves, friendly fenway, cubhouse, thenewyorkyankeesfanclub, ladodgersbestfans, stlouiscardinalsdugout, diehardcubsfans, metsfanzunited, vroc-f2, legendcars, gpl-racing, theofficialnascartodayclub, gpma, cartracing, hockhist, hockeyfansdroppingthegloves, centralpuck, detroitredwingfans, buffalosabresfanatics, sjsharktank, flyersfanatics, pittsburghpenguinfanspage, torontomapleleafsclub

4. Announcements were sent to the following .message boards on leading media and sport sites:

Boston Globe, Chicago Tribune, Dallas News, Denver Post, SFGate, ABC News, CNN, New York Times, CBS News, Nando Times, MSNBC, Yahoo-Daily News, Detroit News, Fox News, NBC, NPR, Cox News, Miami Herald, espn.com, cnnsi.com, sportingnews.com, usatoday.com, nba.com, nhl.com, cbssportsline.com, foxsports.com, yahoo sports, sportsnews.com, MSNBC sports, nfl.com, sportserver.com, mlb.com.

5. We submitted to the following sites: searchenginewatch.com/webmasters, www.esiteblast.com/meta.htm and www.usubmit.com.

References

Alexander, J.E., & Tate, M.A. (1999). *Web wisdom: How to evaluate and create information quality on the We*. Hillsdale, NJ: Lawrence Erlbaum.

Althaus, S.L., & Tewksbury, D. (2000). Patterns of Internet and traditional news media use in a networked community. *Political Communication, 17*, 21-45.

Austin, E.W., & Dong, Q. (1994). Source v. content effects on judgments of news believability. *Journalism Quarterly, 71*, 973-983.

Babbie, E. (2001). *Survey research methods*. Belmont: Wadsworth.

Beard, M. (2002). Fully skinny on sport surfers: Who they are, where they go, and what they buy. [Online] available at www.medialifemagazine.com

Bucy, E.P. (in press) Media credibility reconsidered: Synergy effects between on air and online news. *Journalism & Mass Communication Quarterly*.

Cai, X. (2002, August). *Are computers and traditional media functionally equivalent?* Paper presented at the Association for Education in Journalism and Mass Communication, Miami, FL.

Carter, R.F., & Greenberg, B.S. (1965). Newspapers or television: Which do you believe? *Journalism Quarterly, 42*, 29-34.

Charney, T., & Greenberg, B.S. (2001). Uses and gratifications of the Internet. In C.A. Lin & D. J. Atkin (eds.), *Communication technology and society: Audience adoption and uses of the new media* (pp. 379-407). Cresskill, NJ: Hampton.

Chew, F. (1994). The relationship of information needs to issue relevance and media use. *Journalism Quarterly, 71*, 676-688.

Culbertson, H., & Stempel III, G. (1986). How media use and reliance affect knowledge level. *Communication Research, 13*, 579-602.

Cyber Dialogue (1999). Entertainment draws two-thirds of US users. [Online] Available at www.nua.com

Cyber Dialogue (2001). Online marketers should target sports fans. [Online] Available at www.nua.com.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Eastin, M.S. (2001). Credibility assessments of online health information: The effects of source expertise and knowledge of content. *Journal of Computer-Mediated Communication, 6*, [Online] Available at www.ascusc.org/jcmc.

Eighmey, J. (1997, June). Profiling user responses to commercial Web sites. *Journal of Advertising Research, 37*, 59-66.

eMarketer (2002). Young Americans opt for a variety of sites. [Online] Available at www.nua.com.

Ferguson, D.A., & Perse, E.M. (2000). The World Wide Web as a functional alternative to television. *Journal of Broadcasting & Electronic Media, 44*, 155-174.

Finberg, H.I., & Stone, M.L. (2002). *Digital journalism credibility study.* Washington, D.C.: Online News Association.

Finn, S. (1997). Origins of media exposure: Linking personality traits to TV, radio, print and film use. *Communication Research, 24*, 507-29.

Flanagin, A.J., & Metzger, M.J. (2000). Perceptions of Internet information credibility. *Journalism & Mass Communication Quarterly, 77*, 515-540.

Flanagin, A.J., & Metzger, M.J. (2001). Internet use in the contemporary media environment. *Human Communication Research, 27*, 153-181.

Gantz, W. (1981). The influence of researcher methods on television and newspaper credibility evaluations. *Journal of Broadcasting, 25*, 155-169.

Gaziano, C., & McGrath, K. (1986). Measuring the concept of credibility. *Journalism Quarterly, 63*, 451-462.

Gilster, P. (1997). *Digital literacy.* New York: John Wiley & Sons, Inc.

GomdaWeb. (1998). Perceptions of reliability. [Online] Available at www.stanford.edu/class/ comm217/ reliability/perception.

Greenberg, B.S. (1966). Media use and believability: Some multiple correlates. *Journalism Quarterly, 43*, 665-670, 732.

Greenberg, B.S., & Roloff, M.E. (1974, November). Mass media credibility: Research results and critical issues. *News Research Bulletin of the American Newspapers Publishers Association, No. 6.* Washington, D.C.: ANPA.

Greenberg, B.S. (1974). Gratifications of television viewing and their correlates for British children. In J.G. Blumler and E. Katz (Eds.), *The uses of mass communications: Current perspectives on gratifications research* (pp. 71-92). Beverly Hills, CA: Sage.

Horrigan, J.B., & Rainie, L. (2002). Getting serious online. [Online]. Available at www.pewinternet.org.

Howard, P.E.N., Rainie, L, & Jones, S. (2001). Days and nights on the Internet: The impact of a diffusing technology. [Online]. Available at http://www.pewinternet.org.

Johnson, T.J., & Kaye, B.K. (1998). Cruising is believing? Comparing media and traditional sources on media credibility measures. *Journalism & Mass Communication Quarterly, 75*, 325-340.

Johnson, T.J. & Kaye, B.K. (2000). Using is believing: The influence of reliance on the credibility of online political information among politically interested Internet users. *Journalism & Mass Communication Quarterly*, 77, 865-879.

Johnson, T.J., & Kaye, B.K (2002a). I heard it through the Internet: Factors that determine online credibility among politically interested Internet users. In A. V. Stavros (Ed*.), Advances in Communications and Media Research* (pp. 181-202) New York: Nova Science Publishers.

Johnson, T.J., & Kaye, B.K. (2002b). Webelievability: A path model examining how convenience and reliance predict online credibility. *Journalism & Mass Communication Quarterly, 79,* 619-642.

Johnson, T.J., & Kaye, B.K. (in press). For whom the Web toils: How Internet experience predicts Web reliance and credibility. .*New Jersey Journal of Communication.*

Kang, M.E., & Atkin, D.J.(1999), Exploring the role of media uses and gratifications in multimedia cable adoption. *Telematics & Informatics, 16,* 59-74.

Kaye, B.K. (1998). Uses and gratifications of the World Wide Web: From couch potato to Web potato. *New Jersey Journal of Communication, 6,* 21-40.

Kaye, B. K., & Johnson, T. J. (1999). Taming the cyber frontier: Techniques for improving online surveys. *Social Science Computer Review*, 17 (3), 323-337.

Kaye, B.K., & Johnson, T.J. (2001, August). *A Web for all reasons: Uses and gratifications of the Web for political information.* Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication, Washington, D.C.

Kaye, B. K., Johnson, T. J. (2002). Online and in the know: Uses and gratifications of the Web for political information. *Journal of Broadcasting and Electronic Media, 46,* 54-71.

Kaye, B.K, & Johnson, T. J. (2003). From here to obscurity?: Media substitution *theory* and traditional media in an online world. *Journal of the American Society for Information Science and Technology, 54*, 260-273.

Kim, S.T, Weaver, D., & Willnat, L. (2001). Media reporting and perceived credibility of online polls. *Journalism & Mass Communication Quarterly, 77,* 846-864.

Kiousis, S. (2001). Public trust or mistrust? Perceptions of media credibility in the information age. *Mass Communication & Society.* 4, 381-403.

Kippax, S., & Murray, J.P. (1980). Using the mass media: Need gratification and perceived utility. *Communication Research, 7,* 335-360.

Korgaonkar, P. & Wolin, L. (1999). A multivariate analysis of Web usage. *Journal of Advertising Research, 39,* 53-68.

LaRose, R., Mastro, D., & Eastin, M.S. (2001). Understanding Internet usage: A social cognitive approach to uses and gratifications. *Social Science Computer Review, 19,* 395-413.

Lin, C.A. (2001a). Audience attributes, media supplementation, and likely online service adoption. *Mass Communication & Society, 4,* 19-38.

Lin, C.A. (2001b, August). Online use activity and user gratification expectations. Paper presented to the Annual Meeting Association for Education in Journalism and Mass Communication, Washington, D.C.

Lin, C.A., & Jeffres, L. (1998). Predicting adoption of multimedia cable service. *Journalism Quarterly*, 75, 251-275.

Mashek, J.W., with McGill, L.T., & Powell III, A.C. (1997). *Lethargy '96: How the media covered a listless campaign*. Arlington, VA: The Freedom Forum.

McLeod, J.M., & Becker, L.B. (1974). Testing the validity of gratification measures through political effects analysis. In J.G. Blumler & E. Katz (Eds.), *The uses of mass communications: Current perspectives on gratifications research* (pp. 137-162). Beverly Hills, CA: Sage.

McLeod, J. M., & Becker, L. B. (1981). The uses and gratifications approach. In D. D Nimmo, & K. R. Sanders (Eds.), *Handbook of political communication* (67-99). Beverly Hills, CA: Sage.

Middleberg, D., & Ross, S. S. (2001) *Media in Cyberspace Study*. New York: Middleberg and Associates.

Mulder, R. (1980). Media credibility: A uses-gratifications approach. *Journalism Quarterly 57*, 474-477.

Mulder, R. (1981). A log-linear analysis of media credibility. *Journalism Quarterly, 58*, 635-638.

Nie, N, & Ebring, L. (2000). Study offers early look at home Internet is changing daily life. [Online] Available at www.stanford.edu/dept/news/pr/ 00/000216internet.html.

Online Publishers Association (2002). Online media users more likely to use the same media brand offline, according to Online Publishers Association Survey. [Online]. Available at www.businesswire.com

Palmgreen, P. (1984). Uses and gratifications: A theoretical perspective. In R.N. Bostrom (Ed.), *Communication Yearbook 8* (pp. 61-72). Beverly Hills, CA: Sage Publications, Inc.

Papacharissi, Z., & Rubin, A.M. (2000). Predictors of Internet use. *Journal of Broadcasting and Electronic Media, 44*, 175-196.

Perse, E. M., & Courtright, J. A. (1993). Normative images of communication media: Mass and interpersonal channels in the new media environment. *Human Communication Research*, 19 (4), 485-503.

Petty, R.E., & Cacioppo. J.T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: William C. Brown.

Pew Research Center (2000). Internet sapping broadcast news audience: Investors now go online for quotes, advice. [Online] Available at www.people-press.org/media00rpt.htm.

Pew Research Center (2001). The changing online population: It's more and more like the general population. [Online] Available at www.pewinternet.org/reports.

Pew Research (2003). Daily and overall Internet population. [Online]. Available at www.pewinternet.org/reports.

Potter, W.J. (1998). *Media literacy*. Thousand Oaks, CA: Sage.

Power, J.D. (2001). J.D. Power and Associates Reports: Internet users prefer content over appearance when visiting sports Web sites. [Online] Available at www.jdpa.com/presspass.

Reagan, J., & Zenaty, J. (1979). Local news credibility: Newspaper vs. TV news revisited. *Journalism Quarterly, 56,* 168-172.

Rimmer, T., & Weaver, D. (1987). Different questions, different answers? media use and media credibility. *Journalism Quarterly, 64,* 28-36, 44.

Robinson, M.J., & Kohut, A. (1988). Believability and the press. *Public Opinion Quarterly, 52*, 174-189.

Roper, B. R. (1977). *Changing public attitudes toward television and other mass media, 1959-1976*. New York: Television Information Office.

Rosales, R. (2001). *Expectancy-value theory: Explaining motivations for Internet use among college students*. Unpublished Ph.D. diss, Southern Illinois University at Carbondale.

Rubin, A.M., & Rubin, R. (1982). Older persons' TV viewing patterns and motivations. *Communication Research, 9*, 287-313.

Screen Digest (2000). Sports sites set to win race for profits. [Online] Available at nua.com.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sundar, S.S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly, 76*, 373-386.

Tewksbury, D., & Althaus, S.L. (2000). Differences in knowledge acquisition among readers of the paper and online versions of a national newspaper. *Journalism & Mass Communication Quarterly,* 77, 457-479.

UCLA Internet Report. (2003). Surveying the digital future: Year three. [Online]. Available at www.ccp.ucla.edu.

Vincent, R.C., & Basil, M.D. (1997). College students' news gratifications, media use and current event knowledge. *Journal of Broadcasting & Electronic Media.*

Wanta, W., & Hu, Y-W. (1994). The effects of credibility, reliance, and exposure on media agenda-setting: A path analysis model. *Journalism Quarterly, 71,* 90-98.

Wathen, C.N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology, 53,* 134-144.

Weaver, D.H., & Buddenbaum, J.M (1979, April 20). Newspapers and television: A review of research on uses and effects. *ANPA News Research Report, 19,* Washington, D.C.: ANPA.

Wenner, L.A. (1982). Gratifications sought and obtained in program dependency: A study of network evening news programs and *60 Minutes*. *Communication Research, 9*, 539-560.

Westley, B.H., & Severin, W.J. (1964). Some correlates of media credibility. *Journalism Quarterly, 41,* 325-335.

Witte, J.C., Amoroso, L.M., & Howard, P.E.N (2000). Research methodology: Method and representation in Internet-based survey tools—Mobility, community, and cultural identity in Survey2000. *Social Science Computer Review, 18,* 179-195.

Wu, W., & Weaver, D. (1997). Online democracy or online demagoguery: Public opinion "polls" on the Internet. *Harvard International Journal of Press/Politics*, *2,* 71-86.

Self-esteem, self-affirmation and threats to self-worth:

Testing a motivational explanation for the third-person effect

Patrick C. Meirick
Assistant Professor
Department of Communication
University of Oklahoma

Submitted to the Communication Theory and Methodology Division of the Association for Education in Journalism and Mass Communication for presentation at the 2003 AEJMC Convention, Kansas City, MO.

Direct inquiries to the author at:
Dept. of Communication
University of Oklahoma
610 Elm Ave.
Norman, OK 73019
Phone: 405-325-1574
Fax: 405-325-7625
E-mail: meirick@ou.edu

RUNNING HEAD: Third-person effect

BEST COPY AVAILABLE

Self-esteem, self-affirmation and threats to self-worth: Testing a motivational explanation for the third-person effect

ABSTRACT

The self-enhancement explanation for third-person effects argues that perceiving oneself as resistant to media messages enhances one's self-esteem. The need to self-enhance can be increased by threats to self-worth or reduced by self-affirmation (Steele, 1988). In Study One, third-person effects did not vary by threat condition or by self-esteem, although those high in self-esteem perceived smaller effects on themselves and others. In Study Two, third-person effects were smaller among those whose self-worth was affirmed.

BEST COPY AVAILABLE

The third-person effect is the tendency of people to believe that media messages have greater effects on other people than on themselves (Davison, 1983; for a review, see Perloff, 1999). The effect is quite robust when the messages in question are perceived as undesirable or the effects harmful – pornography, libelous news stories, negative political ads, controversial product ads, etc. – as has usually been the case in third-person research (e.g., Cohen, Mutz, Price and Gunther, 1988; Rucinski and Salmon, 1990; Shah, Faber and Youn, 1999; Thompson, Chaffee and Oshagan, 1990). But for desirable messages, the third-person effect is attenuated (Brosius and Engel, 1996; Duck and Mullin, 1995; Innes and Zeitz, 1988; Gunther and Mundy, 1993) or reversed into a "first-person effect" in which people see themselves as more influenced than others (Chapin, 2000b; Cohen and Davis, 1991; Gunther and Hwa, 1996; Hoorens and Ruiter, 1996; Meirick, 2000; Price, Tewksbury and Huang, 1998).

This message desirability contingency has led some communication researchers to posit a self-enhancement explanation for first- and third-person effects (Gunther and Mundy, 1993; Hoorens and Ruiter, 1996; Perloff, 1999; Meirick, 2000). Self-enhancement is described as "the tendency to see one's own traits, abilities and prospects in a somewhat exaggeratedly rosy light" (Taylor and Brown, 1988), and it is associated with building ourselves up and with tearing others down (if we can justify it) so we look good by comparison (Kunda, 1999). It is the Western self's chief mechanism for maintaining and bolstering self-esteem. It would be consistent with self-enhancement to think ourselves relatively invulnerable to harmful media influence yet relatively receptive to pro-social media messages. Some communication researchers have noted the similarity between the third-person effect and unrealistic optimism (Weinstein, 1980, 1989; L. Perloff and Fetzer, 1986), which pertains to people's belief that compared to others, they are more likely to experience positive events and less likely to experience negative events in the future. Unrealistic optimism, also known as optimistic bias, is generally regarded as motivated by the need for self-enhancement (Taylor and Brown, 1994), and it is presumed that third-person effect is similarly motivated. However, few studies have gone beyond manipulating message desirability in testing the self-enhancement explanation for the third-person effect.

One experimental approach in social psychology may provide the means for a more direct test of the self-enhancement explanation. Steele's work on self-affirmation (Steele and Liu, 1983; Steele, 1988) suggests that people respond to specific threats to self-worth by seizing on opportunities to affirm valued (but sometimes unrelated) aspects of the self. While Steele and Liu (1983) originally addressed such threats in the context of dissonance studies, they and others researchers have applied the self-threat approach to other areas of self-enhancement, such as self-serving theories of success (Dunning, Leuenberger and Sherman, 1995), derogating others (Fein and Spencer, 1997; Branscombe and Wann, 1994), and exaggerating self-ratings of social qualities (Brown and Smart, 1991). In each case, subjects responded to threats to the self-concept by manifesting the various forms of self-enhancement at greater levels than subjects who weren't threatened. Lest anyone argue that these responses do not implicate self-esteem, there is evidence that the response to self-threat interacts with the individual's self-esteem such that people with high self-esteem enhance themselves more than those low in self-esteem do (Brown and Smart, 1991). Moreover, when people are given an opportunity to affirm the self through another activity, self-enhancing biases diminish or disappear.

Steele's approach has become established enough that Fein and Spencer (1997) and Dunning et al. (1995) used self-threat and self-affirmation manipulations to test the motivational underpinnings of outgroup derogation and self-serving theories of success. This approach may provide a similar test for the third-person effect. That is the purpose of the present research.

Study One will manipulate threats to self-worth, which may increase third-person effects; Study Two will manipulate self-affirmation, which should attenuate third-person effects.

Beyond message desirability: Recent third-person effect studies

The self-enhancement explanation for the third-person effect seems to pass the message desirability test. But so far, few other tests have been conducted. In what appears to be the only other experimental examination of the self-enhancement explanation, Brosius and Engel (1996) attempted to make admissions of influence upon the self more palatable in one condition by casting the media effects questions in ways that offered the respondent control over the effect, asking "Do you let yourself be influenced?" versus asking "Do the media influence you?" in the other condition. They expected to find a diminished third-person effect when the respondent could be an active participant in his influence and thereby maintain self-determination, but the differences didn't reach significance. This was an intuitively appealing idea, but the experimental manipulation had no track record. Moreover, while maintaining agency could be enhancing, it's possible that it would be even less enhancing to conceive of oneself as willingly allowing undesirable influence.

A few recent studies have explored correlations between the third-person effect and self-esteem. This is relevant to the self-enhancement explanation because those high in self-esteem are thought to use self-enhancement mechanisms more than those low in self-esteem (Taylor and Brown, 1988). However, the findings have been inconsistent.

David and Johnson (1998) examined the third-person effect of media images of perfect bodies and found a stronger third-person effect for people with high self-esteem. But Chapin (2000a) found a non-significant negative correlation ($r=-.14$) between self-esteem and the perception that others were more affected than oneself. However, it should be noted that the stimuli Chapin used were two safer-sex messages, which many would consider smart to be influenced by, so a negative correlation with the third-person effect (or a positive correlation with the first-person effect) would be consistent with the expectation that those high in self-esteem would have more self-enhancing responses. Banning (2001) hypothesized that self-esteem would have a positive linear relationship with the third-person effect, but found no such relationship. He also expected to find a curvilinear relationship in which those highest in self-esteem would have a smaller third-person effect than those who scored just moderately high; that did not materialize, either.

Finally, there have been two recent studies that examined the third-person effect's correlation with optimistic bias, one form of self-enhancement. Chapin (2000a) examined third-person effects among teenagers in response to safe-sex messages. It also asked teens optimistic bias questions about their relative likelihood of contracting HIV/AIDS. Chapin expected to find a positive correlation between third-person effect and optimistic bias - that teens who thought themselves relatively less likely to be persuaded also would be perceive themselves as less likely to be infected. Instead, he found a small negative relationship. In hindsight, Chapin (2000a) attributed this result to the pro-social nature of the safe-sex messages. Chapin (2000a) argued that optimistic bias would correlate more strongly with the third-person effect if the messages in question were undesirable messages, such as those typically used in third-person effect studies.

Salwen and Dupagne (2000) set out to put that assertion to the test. Toward the end of 1999, they conducted a survey asking people about the likelihood of Y2K problems for themselves and for other people. They also asked them how strongly they agreed or disagreed that news coverage had had a significant impact on their (other people's) beliefs about Y2K. Although Salwen and Dupagne (2000) found evidence of third-person effect and of optimistic bias, they observed almost no correlation (r=.04, ns) between the two. However, news coverage was not necessarily perceived as undesirable. Perceived quality of news coverage had a -.21 beta (p<.001) in predicting third-person effect, a finding is consistent with message quality findings that support a self-enhancement explanation (White, 1997). The fact that message desirability was ambiguous may again explain why there was no correlation between third-person effect and optimistic bias. The negative events described in the optimistic bias portion of the study were hardly ambiguous as to their desirability. Thus, a first-person effect among those who found the news coverage desirable and a third-person effect among those who found it undesirable could both be routes to self-enhancement, but they would look like noise if one was expecting a clear correlation between third-person effect and optimistic bias.

Of course, as Salwen and Dupagne (2000) conceded, "a single-issue study cannot adequately explain the relationship between judgments about experiencing events and judgments of media effects about the events" (p. 22). Moreover, the relationship between optimistic bias and third-person effect is not necessarily relevant to the relationship between self-enhancement and third-person effect because it bypasses the presumed underlying cause: the need to maintain and enhance self-esteem. A more appropriate test might manipulate the extent to which people feel the need to enhance self-esteem. And self-threat and self-affirmation may provide such tests.

Threats to self-worth and self-enhancement

Steele (1988) developed self-affirmation theory to explain how people protect themselves from threats to self-worth. The central idea of self-affirmation theory is that people may counter a threat to one aspect of the self by affirming an unrelated aspect. The means of self-affirmation vary, but these processes and strategies appear to serve largely the same function and can, in fact, be substituted for one another to accomplish the goal of enhancing self-esteem. Tesser (2000; 2001; Tesser and Cornell, 1991) describes this as the "confluence of self-esteem maintenance processes."

For example, Fein and Spencer (1997) argued that stereotyping and prejudice may be a common way for people to maintain their self-image. To test this argument, they manipulated threat to self-worth. Subjects were more likely to evaluate a gay man or a Jewish woman stereotypically or negatively if they had received negative feedback (rather than neutral or positive) on an intelligence test. Among those who had been threatened, derogating an outgroup member mediated an increase in self-esteem. Similarly, other researchers (Sinclair, 1998; Kunda and Sinclair, 1999) have found that white men derogated black and female evaluators when the evaluators evaluated them poorly, but not when they evaluated them highly.[1]

Dunning, Leuenberger and Sherman (1995) applied the self-threat paradigm to Dunning's ongoing study of self-serving theories of success (e.g., Dunning, Perie and Story, 1991; Dunning and Cohen, 1992). This tendency to describe prototypes of excellence in ways that reflect one's own traits appears similar to overly positive views of self, one of the self-enhancement biases identified by Taylor and Brown (1988). Dunning, Leuenberger and Sherman (1995) found that subjects who were told they failed at an intellectual task 1) had self-serving theories about the attributes essential to a successful marriage and 2) evaluated targets similar to themselves more favorably than they did targets unlike themselves. Subjects who were told they had scored highly on the intellectual task showed no such self-enhancement.

---

[1] Note that non-minorities were not derogated, even when self-worth was threatened. Kunda (1990) suggests that negative stereotypes provide justification for derogation. Third-person effect does on some level suggest a derogation of others, but only relative to the self: others are not as sharp as I am. Also, to the extent that the "third persons" can be considered part of an outgroup, derogation would perhaps be justified.

The evidence is quite clear that threats to self-worth lead people to perceive themselves in a self-enhancing light (Brown and Smart, 1991; Dunning et al, 1995). It simply makes sense that self-enhancement can achieved through positive perceptions of the self.

H1: *Compared to those who aren't threatened, people whose self-worth is threatened will perceive undesirable media messages to have less influence on themselves.*

There is some evidence that self-threat leads to less favorable evaluations of others (Fein and Spencer, 1997; Beauregard and Dunning, 1998) or outright sabotage of others' performance (Tesser and Cornell, 1991), but the evidence has qualifiers. Tesser and Cornell's (1991) study involved a special set of conditions: a close other who outperforms the self at a self-relevant task. In other studies where people derogated others after a threat to self-worth, the others who were derogated were members of outgroups; ingroup members weren't derogated (Sinclair, 1998; Kunda and Sinclair, 1999). So, to the extent that self-enhancement can be derived from believing in the gullibility of one's fellow human, we would expect that

H2: Compared to those who aren't threatened, people whose self-worth is threatened will perceive undesirable media messages to have more influence on others.

Given these predictions for perceived effects on self and others, the prediction for the third-person effect becomes clear:

H3: Compared to those who aren't threatened, people whose self-worth is threatened will exhibit a stronger third-person effect; that is, they will perceive undesirable media messages to have less influence on themselves compared to others.

Self-esteem and responses to self-threat

Much of the work on self-enhancement seems predicated on the assumption that these processes are universal. While they are certainly prevalent, at least in Western culture, there are individual differences in the extent to which people employ them. Taylor and Brown (1988) reviewed numerous studies demonstrating that people with high self-esteem tend to actively cope with failure, while those with low self-esteem tend to accept it. Those high in self-esteem are more likely than those low in it to claim more responsibility for success than failure (see Blaine and Crocker, 1993, for a review). However, Campbell (1986) found that while those high and low in self-esteem both take credit for success, those high in self-esteem are more likely to attribute failure to temporary situations, while those low in self-esteem tend to blame themselves (Campbell, 1986).

Campbell's (1986) finding suggests that in dealing with self-esteem, differences come to the fore when dealing with events that threaten self-worth. A number of studies employing the self-threat approach have borne out this differential response. Brown and Smart (1991) found that people with high self-esteem who failed at a purported intelligence test later exaggerated their social (but not intellectual) abilities on a questionnaire and were more likely to help a graduate student. By way of contrast, those with low self-esteem who failed the test denigrated their social skills and were less helpful than low self-esteem subjects who succeeded. This is consistent with other findings that for people high in self-esteem, their personal strengths become highly cognitively accessible in the wake of personal failure, while people with low self-esteem showed no evidence of similar enhancement after failure (Dodgson and Wood, 1998).

These studies suggest that threats to self-worth have greater positive consequences for self-perceptions upon those high in self-esteem than those low in self-esteem. Resistance to undesirable messages is a valued trait and perceiving oneself to have great resistance would be self-enhancing, so it should show the same pattern with regard to threat and self-esteem.

H4: Among those who are threatened, people with high self-esteem will perceive undesirable media messages to have less influence on themselves than will those with low self-esteem.

As discussed earlier, there is some qualified evidence that people may self-enhance by derogating others in response to a self-threat. Some of that evidence also suggests that this response is more likely among those who are relatively high in self-esteem. Beauregard and Dunning (1998) found that subjects' evaluations of another person's intelligence were more negatively related to their own after a threat to self-esteem than after self-affirmation; this difference was greater among those with high self-esteem than those with low self-esteem. Attributing gullibility to people is analogous to attributing stupidity to people, and it appears that those high in self-esteem may be more likely to do so to self-enhance when threatened than those low in self-esteem. Therefore:

H5: Among those who are threatened, people with high self-esteem will perceive undesirable media messages to have more influence on others than will those with low self-esteem.

If we expect those high in self-esteem to perceive less media influence on themselves and more influence on others than do those low in self-esteem, the prediction for the third-person effect is determined:

H6: Among those who are threatened, people with high self-esteem will exhibit a stronger third-person effect - that is, they will perceive negative media messages to have less influence on themselves compared to others – than will those with low self-esteem.

Study 1

Key variables

Threat to self-worth: Threat to self-worth was operationalized with two versions of a purported reasoning test and feedback thereon. To enhance the perceived validity of feedback where success and failure are randomly assigned, McFarlin and Blascovich (1984) created two tests out of the 10 easiest and 10 hardest items from the Remote Associates Test (Mednick, 1962), a word-association task originally designed to test creativity. After taking one of the tests, as in Brown and Smart (1991), subjects were told they'd scored in either the 85th percentile (not threatening to self-worth) or 35th percentile (threatening to self-worth). The manipulation appeared successful. Those given the easy version of the test scored much higher (6.76 out of 10) than those given the hard version (1.42 out of 10, t_{145}= 17.850, $p<.001$). Prior to getting feedback, subjects were asked to rate their performance and their own level of "cognitive flexibility," the trait purportedly being measured. Those given the easy version rated their performance higher (4.39 on a 1-to-7 scale) than those given the hard version (2.29, t_{145}=8.859, $p<.001$), and they rated their ability higher as well (4.85 vs. 3.77, t_{145}=4.343, $p<.001$).

Self-esteem: A seven-point version of the 10-item Rosenberg (1965) self-esteem scale was used to measure global self-esteem. Respondents indicate their level of agreement or disagreement with questions such as "On the whole, I feel satisfied with myself" and "At times I think I am no good at all" (reverse-scored). The scale was highly reliable, with a Cronbach's alpha of .90. Scales scores ranged from 70, the highest possible, to 13. The mean scale score was 55, with a median of 57 and a standard deviation of 9.45. Where required for analysis, the scale was dichotomized and those scoring in the middle quintile were omitted. The low self-esteem group consisted of those who scored 54 or lower, while the high self-esteem group consisted of those scoring 63 or higher.

Third-person effect: An overall scale was created to assess the magnitude of the third-person effect for each subject. The third-person effect was assessed with sets of questions about the impact of different types of undesirable media messages on themselves, others in their class and the public in general. Specifically, subjects were asked about: cigarette advertising effects on attitude toward brand and likelihood to smoke; diet pill infomercial effects on belief in the effectiveness of the product and likelihood to buy; psychic hotline ad effects on belief in the effectiveness of the psychics and likelihood to call; and Sept. 11 news effects on attitudes toward flying and attitudes toward Arabs and Muslims.

The format for the answers was typically a bipolar scale from –3 (much less likely/favorable) to +3 (much more likely/favorable)[2]. The rationale for a –3 to +3 scale was twofold. First, it avoids the potential pitfall of assuming the direction of an effect. Second, it recognizes that people desire autonomy in their attitudes and behaviors. They resist influence through counterarguing and source derogation (Wright, 1973; 1980) such that their attitudes may move in the opposite direction from the persuasive intent of the message (Hoorens and Ruiter, 1996).

To explore the specific hypotheses concerning perceived effects on self and others, scales were created to assess these effects. Alphas were .77 for the 16-item scale of effects on others and .59 for the eight-item scale of effects on self. To address the hypotheses concerning the third-person effect itself, the 16 self-students and self-public difference scores were summed to create a third-person effect scale. The questions about Sept. 11 news were reverse-coded so that positive scores reflected greater message-consistent persuasion (fear of airlines and antipathy toward Arabs). For ease of interpretation, the scale score sums were divided by 16 to show a mean self-other difference on the –3 to +3 scale.[3]

---

[2] Two sets of purchase behavior questions gave response options on a 1 (never) to 7 (definitely) scale. They were worded "After seeing (diet pill infomercials/ psychic hotline ads), how likely do you think it would be that (you/ other students in this class/the public in general) would (buy the pills/call a psychic)?" In calculating the mean scores for the scales, these responses were transformed to a -3 to +3 metric.

Method

Subjects: The 151 subjects were recruited from introductory, intermediate and advanced classes in journalism. The average age was 20.9, 73.5 percent were female, 83.2 percent were white, and 64.9 percent were sophomores or juniors. Initially, a $250 cash drawing was offered as the lone incentive for participation. Recruitment was eventually extended to classes that offered extra credit, which accounted for 85.4 percent of the subjects. Four subjects who scored between 0 and 2 on the easy version of the test (and were therefore unlikely to believe they had scored in the 85th percentile) were removed from analyses for hypothesis testing.

Procedure: Sessions were scheduled over the course of two months. The experimenter greeted the subjects, who arrived one or two at a time, and then took them to individual rooms. He then explained the study, went over the consent forms and gave them the first questionnaire, a 10-item self-esteem measure (Rosenberg, 1965). A few minutes later, the experimenter returned to introduce the "Cognitive Flexibility Test," which was described to them as a measure of reasoning ability that was related to college GPA and scores on standardized tests like the GRE and the LSAT. Subjects were randomly assigned[4] to take an easy or difficult 10-item version of the Remote Associates Test (Mednick, 1962) and allowed five minutes to complete it. After taking the test, subjects were given a feedback form that showed the correct answers, the number of correct answers they had, and their purported percentile rank among college students nationwide. Those given the easy test were told they'd placed in the 85th percentile; those given the hard version were told they'd placed in the 35th percentile. Then, subjects evaluated the impact of different types of media messages on themselves, other students in their class, and the public in general. That questionnaire concluded with demographics questions and a probe for what they study was about. Finally, subjects were debriefed. In all, the procedure took 25 to 30 minutes.

---

[3] Reliability of difference score scales can be calculated as the average reliability of the components minus their correlation, all of which is divided by 1 minus the correlation (Cohen and Cohen, 1983). Using this formula, the reliability of the third-person effect scale is low: .42. Reliability of the self-other difference score is not usually a preoccupation of third-person effect researchers. Typically, the third-person effect is assessed either by examining self-other differences on an item-by-item basis (e.g., Hoorens and Ruiter, 1996) or by examining difference scores of self and other indexes without reporting the reliability of the difference score (e.g., Chapin, 2000a; Shah et al., 1999). However, McLeod et al. (1997) managed to achieve respectable reliabilities (.60 to .74) using Cohen and Cohen's (1983) formula on their three-item scales of effects of rap lyrics on target groups' knowledge, attitudes and behavior. In the present research, the topics vary much more widely, which makes high reliabilities difficult to obtain. But despite the diversity of the scale's topics, it may provide additional power through aggregation. Any statistically significant results found using the overall scale will be in spite of its low reliability, not because of it.

[4] Random assignment in all studies was done using the Research Randomizer at www.randomizer.org, a Web site run by social psychologists Geoffrey C. Urbaniak and Scott Plous. The Randomizer uses a Java applet that taps an algorithm to produce near-random numbers.

Results

The third-person effect: An overall test of the third-person effect showed that the mean perceived effect on others (.56 on a –3 to +3 scale) was greater than the mean perceived effect on self (-.91, t_{150}=26.866, p<.001). Examined individually, effects on others were significantly greater than effects on self for each of the 16 self-other comparisons. All paired samples t's were greater than 6, and all p levels were less than .001.

Hypotheses 1, 2 and 3: The first set of hypotheses predicted that, compared to those who weren't threatened, people whose self-worth was threatened would exhibit smaller perceived effects on self, greater perceived effects on others, and a stronger third-person effect. However, as Table 1 shows, none of these predictions were accurate. Threat condition had no effect on perceived effects on self, contrary to H1 (means of -.91 for nonthreatening vs. -.87 for threatening, t_{145}=.32, ns). Nor did threat affect perceived message effects on others, contrary to H2 (means of .573 for non-threatening vs. .574 for threatening, t_{145}= .01, p>.99).[5] Similarly, there were no differences between the threat conditions in the size of the difference between effects on self and effects on others, so H3 finds no support. In the threat condition, the average effect on others was 1.45 points higher than effects on self; the self-other difference in the non-threatening condition was 1.48, virtually equal, t_{145}=.288, ns.

---

[5] The two groups of others did differ in perceived influence, but threat had no effect on perceived effects for either classmates (t 145=.152, p=.88) or the public (t 145=.165, p=.87).

Hypotheses 4, 5 and 6: This set of hypotheses predicted that, among those who were threatened, people with high self-esteem would exhibit smaller perceived effects on self, greater perceived effects on others, and a stronger third-person effect than would those with low self-esteem. See Table 2. Those high in self-esteem perceived smaller harmful media effects on themselves *and others* than did those low in self-esteem. Zero-order correlations found that among those who were threatened, self-esteem was negatively correlated with both perceived effects on self (r=-.21, p<.10), in marginal support of H4, and on others (r=-.28, p<.05), in contradiction of H5. To better illustrate these relationships (and to eliminate those neither high nor low in self-esteem), t-tests were run comparing high and low self-esteem groups. Among those who were threatened, those high in self-esteem expressed greater resistance to harmful media influence (-1.11) than those low in self-esteem (-.67), t_{57}=2.757, p<.01. Likewise, they perceived smaller effects on others (.40 vs. .81), t_{57}=2.681, p=.01. Among those whose self-worth was not threatened, self-esteem had no significant effect on perceived effects on self, (-.99 for high vs. -.76 for low, t_{56}=1.249, p>.20) or others (.55 for high vs. .59 for low, t_{56}=.253, ns).

As for the third-person effect, there were no differences to be found. The simplest test, a zero-order correlation using only subjects who were threatened (N=73), made use of the full range of the self-esteem scale, but found no relationship (r=-.05, ns). To more clearly illustrate the effect (or lack thereof), t-tests were run using dichotomized self-esteem with the threatened subjects. Those high in self-esteem showed a third-person effect with a mean self-other difference of 1.50, virtually the same as the 1.48 difference for those with relatively low self-esteem (t_{57}=.120, ns). H6 is not supported.

Discussion for Study 1

The third-person effect manifests itself in the absence of self-threat, but it would appear that the third-person effect is not affected by threats to self-worth, nor are perceived effects on self and others. Yet a self-process appears to play a role; self-esteem is negatively related to perceived effects of undesirable messages on both self *and* others. This was the case across experimental conditions, but particularly in the presence of a threat to self-worth, where the differences between self-esteem conditions became statistically significant. Those high in self-esteem did not perceive significantly smaller effects on themselves under threat, nor did those low in self-esteem perceive significantly greater effects on others under threat, but threat amplified the differences between the two groups.

The result here is somewhat consistent with David and Johnson's (1998) finding that self-esteem was negatively correlated with perceived effects on both self and others, although they also found a greater third-person effect for those with high self-esteem, which was not at all the case here. The subject of David and Johnson's (1998) study, pictures of models with unattainable shapes, may have been particularly apt for self-esteem to influence the third-person effect. It may also be that women with high self-esteem are accurate in their perceptions that other women are more likely to be affected by such pictures. This latter interpretation would be consistent with Peiser and Peter's (2001) "limits/possibilities perspective" stating that one's perceptual position, *vis a vis* education or media use, for example, may affect the tendency one has to manifest the third-person effect.

The perception that one has great resistance to harmful messages would be consistent with self-enhancement, and it makes sense that those high in self-esteem would be more likely to manifest this belief. But from a self-enhancement perspective, high self-esteem people perceiving relatively small effects on *others* wouldn't be expected. It might make sense to think of it from the other perspective, that people *low* in self-esteem perceive *greater* effects on others. People with low self-esteem in the threat condition perceived mean effects on others of .81, which post-hoc LSD tests showed was significantly higher than the means for people with high self esteem in both the non-threatening (.55, $p<.10$) and threatening (.40, $p<.01$) conditions. The threatened high-self-esteem group differed significantly only from the threatened low-self-esteem group.

It seems that people with high self-esteem enhance themselves through low estimates of negative effects on self, which can be easily reconciled with their positive self-beliefs. People with low self-esteem, on the other hand, may find it harder to believe in their own efficacy and resistance to persuasion, so enhancement through perceived effects on self is more difficult. Indeed, people tend to prefer consistent self-conceptions (Swann, 1982; Secord and Backman, 1965).

There is some evidence that people with low self-esteem may instead seek self-enhancement indirectly, for instance, themselves by enhancing groups with which they are closely affiliated (Brown, Collins and Schmidt, 1988). Are the higher estimates of negative effects on others on the part of low self-esteem people a similar attempt at indirect self-enhancement? Unlikely. Self-enhancement through derogating others is practiced more by those high in self-esteem (Beauregard and Dunning, 1998).

In any case, it appears likely that self-enhancement processes do play some role in perceptions of media effects. But if so, why didn't a threat to self-worth affect these perceptions? A possible explanation is that there may be a ceiling on people's vigilance against influence from undesirable messages, so that the threat manipulation simply couldn't move them much further in that direction. After all, the third-person effect has been found in dozens of studies that did not experimentally threaten people's self-worth. People don't need any prompting to consider themselves more resistant than others to harmful messages. Furthermore, the messages used in this study were chosen for their high degree of undesirability. They would have been found noxious whether or not one's self-worth was threatened, and that may have contributed to a ceiling effect. When perceived resistance to an undesirable message is already very high, there may be no self-enhancement to be derived from perceiving more resistance. Indeed, extreme "knee-jerk" resistance is a trait that is frowned upon. There just may have been no room for perceived effects on self to decrease, or the third-person effect to increase, in this study.

However, a self-*affirmation* manipulation would not be hampered by a ceiling effect on vigilance. Quite the opposite. Prior self-affirmation would presumably reduce the need to self-enhance through believing oneself resistant to harmful media messages. That is one reason that a self-affirmation study is necessary.

Self-enhancement and self-affirmation

Some of the first work addressing the reduction of need for self-enhancement was done by Steele (Steele and Liu, 1983), who reinterpreted dissonance studies from a self-affirmation perspective. In this view of dissonance reduction, an individual's positive self-image is threatened by the awareness of having done something foolish or wrong, and an attitude change can make the errant behavior seem more reasonable, thus diminishing the blow to self-worth. Steele and Liu (1983) made their case for this view of dissonance reduction by providing another path to self-affirmation – calling to mind valued aspects of the self – for some subjects who had been induced to write an essay favoring a policy they did not support. In one set of conditions, subjects were given a questionnaire about politics and economics prior to the attitude questionnaire. For those subjects who cared about politics and economics, attitude change was minimal; they had already enhanced themselves by answering the questionnaire.

Further support for the notion of self-affirmation through alternate paths came within the context of Tesser's self-evaluation maintenance model (Tesser, 1980; Tesser and Campbell, 1983; Tesser and Cornell, 1991), which Tesser was putting forward as Steele was formulating his self-affirmation model. According to Tesser's model, people will act to minimize the threat to self-worth posed by someone else's superior performance by becoming less close to that person, reducing the self-relevance of the performance domain or minimizing the other's performance. In Tesser and Cornell (1991), when subjects believed that a word game was related to intelligence (i.e., that it was self-relevant) – and that a friend and a stranger had outperformed them in the first round (i.e., self-worth was threatened) – subjects aided the stranger rather than the friend. They helped the friend when they believed the game was just a game, presumably to bask in the friend's reflected glory. But both patterns of differential helping were eliminated when another opportunity for self-affirmation was provided.

Prior self-affirmation may also serve to pre-empt self-enhancing biases in the absence of a threat to the self. Fein and Spencer's (1997) work on prejudice also included a study in which half of the subjects had an opportunity for self-affirmation via writing a few paragraphs about a value important to them, and half did not. The self-affirmed subjects evaluated a Jewish and a non-Jewish job candidate equally favorably, while those who were not self-affirmed evaluated the Jewish candidate more negatively.[6] Positive reinforcement through a high score on an intellectual task kept people in the theories-of-success study by Dunning et al. (1995) from concocting self-enhancing theories of success. This came in lieu of a threat and seemed to have served the same purpose as the prior self-affirmation described here. Moreover, Dunning has documented self-serving theories of success in the absence of self-threats; people employ them spontaneously, as they do with other self-enhancing mechanisms like optimistic bias (Weinstein, 1980; 1989) and self-serving attribution bias (Snyder, Stephan and Rosenfield, 1976). Yet getting prior affirmation apparently obviated the need to self-enhance. This suggests that people do not self-enhance exclusively to heal a wounded ego, nor do they self-enhance at every possible opportunity. Apparently threat is not a necessary prerequisite for self-enhancing behavior; in some cases, it may merely whet an already healthy appetite for self-enhancement. But it also seems that people's appetite for self-enhancement has limits, that they can be temporarily sated with self-affirmation so that they pass on seconds, so to speak.

The results of Dunning et al. (1995) suggest that when the need to self-enhance is reduced, perceptions of the self won't be as exaggeratedly positive. It follows that a similar intervention may leave people less in need of viewing themselves as invulnerable to media influence.

H7: Compared to those who aren't given an opportunity to self-affirm, people who self-affirm will perceive undesirable messages to have greater effects on themselves.

---

[6] Fein and Spencer (1997), citing Spencer and Steele (1990), are careful to note that the self-affirmation manipulation they used, "in the absence of self-image threat, ... does not affect participants' state self-esteem" (p. 32).

There is evidence that self-enhancement via outgroup derogation (Fein and Spencer, 1997) or sabotage of rivals (Tesser and Cornell, 1991) is lessened when an opportunity to self-affirm is provided first. This suggests that the need to perceive others as vulnerable to media influence might be smaller for those who self-affirm.

H8: Compared to those who aren't given an opportunity to self-affirm, people who self-affirm will perceive undesirable messages to have smaller effects on others.

To the extent that the previous two hypotheses are upheld, we would also expect that:

H9: Compared to those who aren't given an opportunity to self-affirm, people who self-affirm will exhibit a weaker third-person effect.

Study 1

Key variables

Effects on self and effects on others: The scales were calculated as described in Study 1. Alphas were .82 for undesirable messages on others and .62 for undesirable messages on self.

Third-person effect: The scale was calculated as in Study 1. Using Cohen and Cohen's (1983) formula for difference-score scales, reliability was .52 for the third-person effect scale.

Self-affirmation: Randomly selected students were given the opportunity to self-affirm using a procedure devised by Fein and Spencer (1997). In the self-affirmation condition, subjects were asked to circle one of four values from the Allport-Vernon values inventory that meant the most to them, then write a few paragraphs about why it was important to them. In another version, subjects were asked to write about the value that was least important to them and write about why it might be important to someone else.

This mode of self-affirmation is a variation on the procedure used by Steele (e.g., Steele and Liu, 1983; Steele, 1988), which had all students write on a given value, which would be important to some and not to others. Fein and Spencer's (1997) procedure ensures that assignment to the affirmation conditions is random and not potentially confounded with a given value.

Method

Subjects: Subjects (N=75) were recruited from an introductory psychology class. The subjects had a mean age of 19.2, 58.1 percent were female, 82.4 percent were white and 66.2 percent were first-year students. Subjects were given credit toward a research experience class requirement. One subject in the self-affirmation condition failed to write anything and was excluded from the analysis.

Procedure: Data were gathered in two large group sessions after class, once in mid-March and once in early May. The two versions of the questionnaire were placed in random order in advance. Subjects were given the questionnaire to fill out in their seats. The first section asked them to circle the value (business/economics, art/music/theater, social life/relationships or science/pursuit of knowledge) that was most important (self-affirmation condition) or least important (no affirmation) to them. In the remaining space on the page, subjects were asked to explain why the value they circled was important to them (self-affirmation) or might be important to someone else (no affirmation). After that, the subjects all filled out the message evaluation questionnaire and demographic questions that were used in Study 1. They then filled out the paperwork required to receive participation credit and were debriefed. Most people finished the questionnaire in 15 minutes.

Results

The third-person effect: An overall third-person effect was found. The mean perceived message-consistent effect of undesirable messages on self was -.84 on a –3 to +3 scale, much lower than the mean of .66 for others. The resulting difference, 1.51, was significant (t_{73}=16.236, p<.001).

Looking at each of the 16 self-students and self-public comparisons individually, all showed effects on others significantly greater than effects on self, with all t's greater than 4.9 and all p levels less than .001.

Hypotheses 7, 8 and 9: It was expected that, compared to those who weren't given an opportunity to self-affirm, people who self-affirm would exhibit greater message-consistent effect on self, smaller effects on others, and a weaker third-person effect. See Table 3. It appeared that those who self-affirmed may have perceived themselves as less resistant to undesirable messages (-.67) than those who did not (-1.00), a difference that approached significance (t_{72}=1.947, p=.055), a qualified confirmation of H7. However, there was no significant difference in perceived effects on others between those who self-affirmed (.61) and those who did not (.72), (t_{72}=.61, ns), so H8 finds no support.

While subjects in both conditions believed undesirable media messages would affect others more than themselves, the self-other gap was indeed smaller for those who did self-affirm (1.28) than those who did not (1.72), a significant difference (t_{72}=2.412, p<.05) in support of H9.

Discussion for Study 2

People who self-affirmed had a smaller third-person effect than those who didn't, which supports a self-enhancement explanation. With their recent self-affirmation, these subjects perhaps didn't feel as great a need to believe themselves relatively invulnerable to harmful media influence. The difference appeared to stem mainly from perceived effects on the self, which is consistent with the findings of Dunning et al. (1995) and the evidence that self-perceptions are usually the conduit for self-enhancement (Brown and Smart, 1991). This result also supports the notion that while self-enhancement need not be triggered by a threat to self-worth, the need for it is finite. If people had an inexhaustible need to enhance themselves, recent self-affirmation would not affect subsequent opportunities to self-enhance. Here, it did.

The findings in Study Two also lend support to the "ceiling effect" explanation for the non-findings in Study One. A threat to self-worth may well have fueled the need to self-enhance. But people already are quite resistant to undesirable messages, and perceiving oneself as rabidly resistant may not have not have presented a viable avenue toward self-enhancement. But Study Two would have avoided any such ceiling effect because the self-affirmation manipulation tends to remove some of the usual pressure to self-enhance. People who self-affirmed perceived themselves to be only moderately resistant to undesirable messages, rather than strongly resistant like the unaffirmed in Study Two (and both groups in Study One).

Overall discussion

Based on the work of Steele (1988) and others who have employed his methodology, it was expected that changing the need for self-enhancement would change the size of the third-person effect in Studies One and Two. To the extent that the third-person effect is driven by self-enhancement, increasing or decreasing the extent to which people seek enhancement was expected to increase or decrease the degree to which people perceive themselves as resistant (and others as vulnerable) to undesirable media messages. As it turned out, this expectation was half-right. Threat to self-worth in Study One had no effect on the size of the third-person effect, but the third-person effect in Study Two was indeed smaller among people who had the opportunity to affirm themselves than among those who didn't. The latter result supports the self-enhancement explanation for the third-person effect. As discussed above, a ceiling on the third-person effect (or more to the point, on people's willingness to perceive themselves as resistant) might explain why changing the need for self-enhancement changed the size of the third-person effect under self-affirmation but not under threat.

Based on research showing that people high in self-esteem self-enhance more than those low in self-esteem, particularly under threats to their self-worth, it had been expected in Study One that people high in self-esteem would demonstrate a larger third-person effect when threatened than would people low in self-esteem. In fact, there was no difference in the size of the self-other gap between the two groups. Those high in self-esteem did perceive themselves to be more resistant to undesirable messages than did those low in self-esteem, but the highs also perceived smaller effects on others than did the lows. These differences between self-esteem groups were significant for those who were threatened but not those who weren't threatened. So in a condition in which we had reason to expect greater self-enhancement among those high in self-esteem than among those with low self-esteem, the only difference that was consistent with the expectation was that highs perceived smaller media effects on themselves than did lows. This suggests that perceptions of effects on oneself may be a more important source of self-enhancement than perceived effects on others or the self-other gap – the third-person effect -- itself. Supporting this suggestion is the fact that where significant differences in the third-person effect were found in between the affirmed and unaffirmed in Study Two, the differences were due almost entirely to differences in perceived effects on self. These findings are consistent with evidence that perceived effects on self are quite affected by manipulations of message desirability, while evaluations of others appear less affected (Gunther and Mundy, 1993; Salwen and Dupagne, 2000; White, 1997). In short, the case for the self-enhancement as an explanation for perceived effects on self finds support in the current research.

However, the already shaky case for self-enhancement's role in perceived effects on others finds no further evidence here. As noted before, derogating others seems to be a relatively rare road to self-enhancement, a road that is followed only if it has been paved with negative stereotypes about the others being derogated (Kunda, 1999). In this research, manipulating subjects' need for self-enhancement had no effect on how they viewed the susceptibility of fellow students or people in general to harmful media messages. It may well be that different results could be found for socially stigmatized "others."

The current research suggests that self-enhancement, through its influence on perceived effects on self, does have a role to play in the third-person effect. This is not to rule out other explanations, especially ones that bear on perceived effects on others. The relative roles of effects on others and effects on self have long been debated in third-person effects research. Perloff's (1993) first review of third-person effect research offered overestimation of media effects on others and/or underestimation of media effects on self as the two explanations for the third-person effect. In studies that compare attitude change in the third-person study with those in a control group, evidence has been found for both explanations, but more consistently for overestimation. Some studies have found support only for overestimation (Gunther, 1991; Lasorsa, 1989; Price et al., 1998), while others have found both overestimation of effects on others and underestimation of effects on self (Cohen et al., 1988; Gunther and Thorson, 1992). The current findings suggest that self-enhancement is a reason that people underestimate media effects on themselves, and that lessening the need for self-enhancement can reduce the amount of underestimation and thus the size of the third-person effect.

This research has a number of limitations. It relies on a student convenience sample, which could inflate the size of the third-person effect (Perloff, 1999) and limit generalizability.

Also, the reliability of the third-person effect difference-score scale was rather low. But these limitations are not crippling. The study was not attempting to measure population parameters in the size of the third-person effect, in which case a student sample would be suspect, but rather the effects of manipulating basic psychological processes. There is no reason to believe that different results would have been found in a sample of the general public. As for the difference score scale, a more reliable scale would have been preferable, but some significant findings were obtained despite the subpar reliabilities.

This is perhaps the first study of the third-person effect in which subjects' need to self-enhance was experimentally manipulated, a procedure that confirms the findings of years of message desirability studies and offers a level of causal inference not permitted by correlational studies. We can say with more confidence than ever before that motivational processes play a role in the third-person effect.

References

Banning, S.A. (2001). Do you see what I see? Third-person effects on public communication through self-esteem, social stigma and product use. Mass Communication and Society, 4 (2), 127-147.

Beauregard, K.S. and Dunning, D. (1998). Turning up the contrast: Self-enhancement motives prompt egocentric contrast effects in social judgments. Journal of Personality and Social Psychology, 74 (3), 606-621.

Blaine, B. and Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. In R. Baumeister (Ed.), Self-esteem: The puzzle of low self-regard (pp. 55-85). New York: Plenum.

Branscombe, N.R. and Wann, D.L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. European Journal of Social Psychology, 24, 641-657.

Brosius, H. and Engel, D. (1996). The causes of third person effects: Unrealistic optimism, impersonal impact, or negative attitudes toward media influence. International Journal of Public Opinion Research, 8, 142-162.

Brown, J.D., Collins, R.L. and Schmidt, G.W. (1988). Self-esteem and direct versus indirect forms of self-enhancement. Journal of Personality and Social Psychology, 55, 445-453.

Brown, J.D. and Smart, S.A. (1991). The self and social conduct: Linking self-representations to prosocial behavior. Journal of Personality and Social Psychology, 60, 368-375.

Campbell, J.D. (1986). Similarity and uniqueness: The effects of attribute type, relevance and individual differences in self-esteem and depression. Journal of Personality and Social Psychology, 50, 281-294.

Chapin, J.R. (2000a) Third-person perception and optimistic bias among urban minority at-risk youth. Communication Research, 27, 51-81.

Chapin, J.R. (2000b). Not like me: Self vs. other distinctions in first-person perception. Communication Research Reports, 17, 320-330.

Cohen, J. and Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. and Davis, R.G. (1991). Third-person effects and the differential impact in negative political advertising. Journalism Quarterly, 68, 680-688.

Cohen, J., Mutz, D., Price, V. and Gunther, A. (1988). Perceived impact of defamation: An experiment on third person effects. Public Opinion Quarterly, 52, 161-173.

David, P. and Johnson, M.A. (1998). The role of self in third-person effects about body image. Journal of Communication, 48 (4), 37-58.

Davison, W.P. (1983). The third-person effect in communication. Public Opinion Quarterly, 47, 1-15.

Dodgson, P.G. and Wood, J.V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. Journal of Personality and Social Psychology, 75, 178-197.

Duck, J.M. and Mullin, B. (1995). The perceived impact of the mass media: Reconsidering the third-person effect. European Journal of Social Psychology, 25, 77-93.

Dunning, D. and Cohen, G.L. (1992). Egocentric definitions of traits and abilities in social judgment. Journal of Personality and Social Psychology, 63, 341-355.

Dunning, D., Leuenberger, A. and Sherman, D.A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? Journal of Personality and Social Psychology, 69, 58-68.

Dunning, D., Perie, M., and Story, A.L. (1991). Self-serving prototypes of social categories. Journal of Personality and Social Psychology, 61, 957-968.

Fein, S. and Spencer, S.J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. Journal of Personality and Social Psychology, 73, 31-44.

Gunther, A.C. (1991). What we think others think: Cause and consequence in the third person effect. Communication Research, 18, 355-372.

Gunther, A.C. and Hwa, A.P. (1996). Public perceptions of television influence and opinions about censorship in Singapore. International Journal of Public Opinion Research, 8, 248-265.

Gunther, A.C. and Mundy, P. (1993). Biased optimism and the third-person effect. Journalism Quarterly, 70, 1, 58-67.

Gunther, A.C. and Thorson, E. (1992). Perceived persuasive effects of product commercials and public service announcements: Third-person effects in new domains. Communication Research, 18 (3), 355-372.

Hoorens, V. and Ruiter, S. (1996). The optimal impact phenomenon: beyond the third person effect. European Journal of Social Psychology, 26, 599-610.

Innes, J.M. and Zeitz, H. (1988). The public's view of the impact of the mass media: A test of the 'third person' effect. European Journal of Social Psychology, 18, 457-463.

Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108, 480-498.

Kunda, Z. (1999). Social cognition: Making sense of people. Cambridge, Mass.: MIT Press.

Kunda, Z. and Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application and inhibition. Psychological Inquiry, 10, 12-22.

Lasorsa, D.L. (1989). Real and perceived effects of "Amerika." Journalism Quarterly, 66, 373-378, 529.

McFarlin, D.B. and Blascovich, J. (1984). On the Remote Associates Test as an alternative to illusory performance feedback: A methodological note. Basic and Applied Social Psychology, 5, 223-229.

Mednick, S.A. (1962). The associative basis of the creative process. Psychological Review, 69, 220-232.

Meirick, P. (2000). Political distance and message desirability: Three studies of political advertising and the third-person effect. Presented at the Association for Education in Journalism and Mass Communication annual convention, Phoenix, AZ, August 2000.

Peiser, W. and Peter, J. (2001). Explaining individual differences in third-person perception: A limits/possibilities perspective. Communication Research, 28 (2), 156-180.

Perloff, L.S. and Fetzer, B.K. (1986). Self-other judgments and perceived vulnerability to victimization. Journal of Personality and Social Psychology, 50, 502-510.

Perloff, R.M. (1993). Third-person effect research 1983-1992: A review and synthesis. International Journal of Public Opinion Research, 5, 167-184.

Perloff, R.M. (1999). The third-person effect: A critical review and synthesis. Media Psychology, 1, 353-378.

Price, V. Tewksbury, D. and Huang, L-N. (1998). Third-person effects on publication of a Holocaust denial advertisement. Journal of Communication, 48 (2), 3-26.

Rucinski, D. and Salmon, C.T. (1990). The "other" as the vulnerable voter: A study of the third-person effect in the 1988 campaign. International Journal of Public Opinion Research, 2, 345-368.

Salwen, M.B. and Dupagne, M. (2000). Optimistic bias and the third-person effect: Public estimations of Y2K effects on self and others. Presented at the Association for Education in Journalism and Mass Communication annual convention, Phoenix, AZ, August 2000.

Secord, P.F. and Backman, C.W. (1965). An interpersonal approach to personality. In B. Maher (Ed.), Progress in experimental personality research (Vol. 2). New York: Academic Press.

Shah, D.V., Faber, R.J., and Youn, S. (1999). Susceptibility and severity: Perceptual dimensions underlying the third-person effect. Communication Research, 26, 240-267.

Sinclair, L. (1998). Justifying desired impressions of evaluators: Motivated activation, application, and inhibition of stereotypes. Unpublished doctoral dissertation, University of Waterloo.

Snyder, M.L., Stephan, W.G. and Rosenfield, D. (1976). Egotism and attribution. Journal of Personality and Social Psychology, 33, 435-441.

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.) Advances in experimental social psychology (Vol. 21, pp. 261-302). New York: Academic Press.

Steele, C.M. and Liu, T.J. (1983). Dissonance processes as self-affirmation. Journal of Personality and Social Psychology, 45, 5-19.

Swann, W.B. (1982). Self-verification: Bringing social reality into harmony with the self. In J. Sols and A.G. Greenwald (Eds.), Psychological perspectives on the self, (Vol. 2), Hillsdale, NJ: Lawrence Erlbaum.

Taylor, S.E. and Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin, 103, 193-210.

Taylor, S.E. and Brown, J.D. (1994). "Illusion" of mental health does not explain positive illusions. American Psychologist, 49, 972-973.

Tesser, A. (1980). Self-esteem maintenance in family dynamics. Journal of Personality and Social Psychology, 39, 77-91.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. Personality and Social Psychology Review, 4 (4), 290-299.

Tesser, A. (2001). On the plasticity of self-defense. Current Directions in Psychological Science, 10 (2), 66-69.

Tesser, A. and Campbell, J. (1983). Self-definition and self-evaluation maintenance. In J. Suls and A. Greenwald (Eds.) Social psychological perspectives on the self (pp. 1-31). Hillsdale, NJ: Erlbaum.

Tesser, A. and Cornell, D.P. (1991). On the confluence of self processes. Journal of Experimental Social Psychology, 27, 501-526.

Thompson, M.E., Chaffee, S.H., and Oshagan, H.H. (1990). Regulating pornography: A public dilemma. Journal of Communication, 40, 73-83.

Weinstein, N.D. (1980). Unrealistic optimism about future life events. Journal of Personality and Social Psychology, 39, 806-820.

Weinstein, N.D. (1989). Optimistic biases about personal risks. Science, Dec. 8, 1989, 1232-1233.

White, H.A. (1997). Considering interacting factors in the third-person effect: Argument strength and social distance. Journalism and Mass Communication Quarterly, 74, 3, 557-564.

Wright, P. (1973). The cognitive processes mediating acceptance of advertising. Journal of Marketing Research, 10, (Feb.), 53-62.

Wright, P. (1980). Message evoked thoughts: Persuasion research using thought verbalizations. Journal of Consumer Research, 7 (Sept.) 151-175.

Wylie, R. (1979). The self concept. Lincoln: University of Nebraska Press.

Tables

Table 1: Perceived effects of undesirable messages by threat condition

	Threat	No threat	t value
Self	-.87	-.91	.320
Others	.57	.57	.010
Third-person effect	1.45	1.48	.288

Note: ***p<.001 **p<.01 *p<.05 ^p<.10

N=73 for threat. N=74 for no threat.

Table 2: Perceived effects of undesirable messages under threat by self-esteem

	Low self-esteem	High self-esteem	t value
Self	-.67	-1.11	2.757**
Others	.81	.40	2.681**
Third-person effect	1.48	1.50	.120

Note: ***p<.001 **p<.01 *p<.05 ^p<.10

N=31 for low self-esteem. N=28 for high self-esteem.

Table 3: Perceived effects of undesirable messages by affirmation condition

	Not affirmed	Self-affirmed	t value
Self	-1.00	-.67	1.947^
Others	.72	.61	.610
Third-person effect	1.72	1.28	2.412*

Note: ***p<.001 **p<.01 *p<.05 ^p<.10

N=38 for not affirmed. N=36 for self-affirmed.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore„does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (1/2003)